

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

Rotating radio waves point to
extreme magnetic environment
for source of repeating fast
radio bursts **PAGE 182**

TWISTED VISTA

NEUROSCIENCE

TOTAL RECALL

*The quest to map memories
in the brain*

PAGE 146

BIOTECHNOLOGY

HALTING HEARING LOSS

*Gene editing in mice
prevents inherited deafness*

PAGES 162 & 217

MATERIALS SCIENCE

EXCITONS TURN ON THE LIGHT

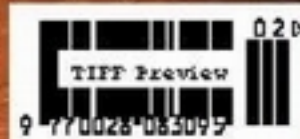
*Bright triplet emission from
perovskite nanocrystals*

PAGES 163 & 189

NATURE.COM/NATURE

11 January 2018 £10

Vol 553, No 7687



THIS WEEK

EDITORIALS

DENIM Chemistry offers a greener route to blue jeans **p.128**

WORLD VIEW Research funders should push for open citations **p.129**



COLOURS Secrets of spider rainbow display revealed **p.131**

Under new management

The mysterious benefactors who are about to take over the Arecibo radio telescope have an early success to celebrate — whoever they are.

When the US National Science Foundation (NSF) drew up a plan to demolish its radio telescope near Arecibo, Puerto Rico, it did conclude that something positive would result — although it was only a minor and short-term benefit. Five specialists in explosives would need to spend a month on the Caribbean island, and, the NSF said in an environmental-impact statement last year, the local community could profit from what the visitors would spend on meals and lodging.

Hoteliers and restaurant owners aside, most of the local workers and researchers who help to keep the giant dish functioning breathed a sigh of relief last November, when the NSF announced that the telescope would remain standing. At least one partner organization had pledged to help fund it, solving a cash crunch at the decades-old facility.

The identity of the saviours is still a closely guarded secret (although everyone in the astronomy community has their own idea of the funders' identity, ranging from overseas agencies to universities). Whoever they are, they are sure to be smiling to themselves this week. Their new toy has shown what it can still do.

In a paper in *Nature* this week, astronomer Daniele Michilli of the University of Amsterdam and his colleagues describe how they used the Arecibo dish to track a mysterious signal from deep space called a fast radio burst (D. Michilli *et al.* *Nature* **553**, 182–185; 2018). These powerful but short-lived flashes of radio noise were first discovered a decade ago, but their source remains unknown. They are one of the biggest outstanding astrophysical mysteries today.

Most of these sources blaze into life just once and then vanish. But a fast radio burst in the constellation Auriga, first spotted in November 2012, has shown itself many times since. Indeed, Michilli and his team recorded at least 16 separate flashes of its activity. Each time, they gleaned a little more information about its probable origin.

The trick, it turns out, lies in looking at the polarization of radiation coming from the burst. The plane of polarization rotates when the light travels through a magnetic field, an effect first seen by physicist Michael Faraday in 1845. For the Auriga burst, the Faraday rotation is large and variable — suggesting that the light must be travelling through a highly magnetized environment.

Until now, this type of Faraday rotation has been seen only close to black holes. So one possible explanation for this fast radio burst is that something is producing radio emissions very near to a black hole. Imagine, perhaps, a dense neutron star burping out radiation that twists and rotates as it travels through its highly magnetized surroundings. The work is the most precise look yet at what could be powering fast radio bursts (or at least one of them).

The announcement of the discovery comes after a tumultuous couple of years for the Arecibo telescope. Alongside the uncertainty over its funding, the facility — like much of Puerto Rico — was battered and put temporarily out of action by Hurricane Maria last year. On restarting its science observations last November, the first

thing the big dish did was to return its gaze to Auriga.

Like many veteran science experiments, Arecibo has an impressive back catalogue. In cinema history, it's where Jodie Foster listened for aliens in 1997's *Contact*, and where Pierce Brosnan's James Bond dispatched villain Sean Bean in *GoldenEye* (1995). In scientific history, the telescope beamed a message meant for extraterrestrials to the globular star cluster M13 in 1974, and has probed dangerous near-

Earth asteroids to help protect the planet from cosmic impacts.

“The discovery comes after a tumultuous couple of years for the Arecibo telescope.”

Now the NSF wants to free up money for newer astronomical facilities by offloading some of its older ones, including Arecibo. With the demolition plan nixed, the current funding arrangement will end in April and the NSF will officially hand the controls to

the mystery newcomers, who have agreed to step in as the agency scales down its annual contributions from US\$8 million to \$2 million over the next 5 years. (NASA will continue to pay one-third of the observatory's costs.)

The dish that the benefactors get for their money is no longer the world's biggest telescope of its type. China switched on its larger Five-hundred-meter Aperture Spherical radio Telescope (FAST) in 2016, and the facility is already making headlines by chalking up discoveries — three new pulsars last month alone. But the sky is a big place, and there is plenty of science to go around. Arecibo is rightly safe from the dynamite for now. ■

Science at sea

Debate on a United Nations treaty to protect the open ocean offers an opportunity for scientists.

In a rare diplomatic breakthrough — and good news for marine scientists and conservationists around the world — nations agreed in 2016 to protect a huge area of ocean off the coast of Antarctica from commercial fishing and other harmful activities. That success came only after years of failed discussions. It was followed by another positive step: in December, Arctic Council countries decided not to fish industrially in the Arctic Sea.

These are good signs. Still missing, though, is a more significant agreement — a mechanism that would allow governments to create marine reserves in ecologically crucial ocean regions beyond any national jurisdiction.

Could the United Nations Convention on the Law of the Sea

fill the gap? The UN thinks so. On 24 December, it convened an intergovernmental conference to produce a legally binding treaty on the conservation and sustainable use of biological diversity in the high seas outside national maritime boundaries. It's a crucial first step, and is encouraging because it suggests that political will is building to draft international rules that protect the ocean wilderness.

The vote, after almost a decade of preparatory work, reflects scientists' growing concern about the alarming state of the global oceans. And public awareness about issues such as overfishing, plastic pollution and species extinction is sharply on the rise in many countries.

The planned treaty, due by 2020, is much needed. A global commons, the high seas cover half of Earth's surface and provide eco-services of immeasurable value. Still, any new pact cannot address all the ills of the seven seas. The surge in plastic waste, for example, has to be tackled at its terrestrial source, mainly with the producers. But a well-crafted and properly enforced rulebook can do much to protect ocean ecosystems from man-made harm.

Any future network of high-seas reserves will need to cover a large variety of species in representative ecosystems in all climate zones. To do this, researchers with the UN Convention on Biological Diversity suggest that marine protected areas should cover at least 10% of the global ocean by 2020. At present, the figure is closer to 6% — almost all in coastal waters. The higher target will be impossible to reach without setting aside reserves in high-seas regions that are as yet legally out of reach. Hence the need for a new treaty.

The treaty's range and scope are yet to be defined, and science has the chance to help frame its demands, and to ensure that the goals of protection and conservation are effectively met. Our understanding of marine ecosystems is best for coastal and inshore regions. An evidence-based approach to protecting the wilderness of the high seas will require massive amounts of research. For example, to get a better sense of the scale of the looming ocean crisis, scientists need to map ecosystem structures and deep-seabed habitats, and to track migratory

patterns of critical species. They will also want to take a closer look at how biological processes in the deep ocean control key chemical cycles, such as carbon uptake and release, that govern Earth's climate.

Research can benefit from, as well as inform, protection. Recent studies show that marine reserves can help species adapt to ocean acidification and other impacts of climate change (see, for example, C. M. Roberts *et al. Proc. Natl Acad. Sci. USA* **114**, 6167–6175; 2017). Such areas can

“Implementation of any rules will have to rely on effective satellite surveillance.”

serve as a control by which to evaluate the impact of fishing and environmental change on marine ecosystems. Researchers can also help to set priorities, by working to identify key ecosystems that need protection from overfishing and other human interference.

A meaningful high-seas pact must also encourage effective fisheries management outside protected areas, to support sustainable catch. And implementation of any rules will have to rely on effective satellite surveillance of fisheries activities on the open ocean. The International Maritime Organization (and Interpol) is already using vessel-monitoring technology to track ship movements and suspicious activity.

The next step will be the first session of the intergovernmental conference, on 4–17 September. It is unclear whether key fishing nations — including the United States, Russia and China — will ratify any agreement. Encouragingly, these countries have not blocked the work of the preparatory committee. Other nations, including Norway, Iceland, Japan and South Korea, have signalled full support for a legally binding instrument. The number of signatures required for the treaty to be enforced is yet to be negotiated.

Whatever arrangements emerge, the UN's move should provide ample research opportunities. Funders should take note. A treaty involving an international research mandate — including a regime to regulate controversial geoengineering experiments such as ocean iron fertilization — would be a boon for ocean health and responsible science. ■

In the jeans

An environmentally friendly way to dye denim could usher in a long-overdue new fashion.

In his Latin description of the Gallic Wars, Julius Caesar wrote: “*se Britanni vitro inficiunt*” — widely translated as meaning that the Britons dyed themselves with woad. Hence, many sources will tell you, the Romans named the ancient people of northern Britain the Picts, or ‘painted ones’. Among the objections to this claim is that woad is not a very good dye for people — it's caustic and irritates the skin and eyes.

It's not a great dye for textiles, either. The indigo colour squeezed from plants including woad (*Isatis tinctoria*) doesn't dissolve in water and so can't penetrate and bind cloth fibres. Instead, it must be chemically converted into a water-soluble compound called leucoindigo, or white indigo, which then adsorbs to the textile surface. It is most commonly used on denim. Over 4 billion denim garments are produced each year. These days, most are dyed blue with synthetic indigo, but the artificial colour must still be fixed using a potent bleaching agent. This is one reason why indigo dyeing is so polluting, as shown vividly by the numerous rivers in China and elsewhere that have been turned blue by untreated waste from jeans factories. According to environmental groups, textile dyeing is one of the most polluting industries in the world.

Indigo dyeing is so widespread that it is hard to replace with a cleaner process. But scientists are trying. Writing online in *Nature Chemical Biology*, researchers describe a more environmentally friendly method of making and applying indigo dye that relies on genetically engineered bacteria (T. M. Hsu *et al. Nature Chem. Biol.*

<http://dx.doi.org/10.1038/nchembio.2552>; 2018).

The process borrows a chemical switch from nature. Inside plant leaves, the unstable indigo precursor indoxyl is combined with glucose and stored as a colourless molecule called indican. The researchers mimicked this by adding genes to *Escherichia coli* bacteria to make them secrete indican. To dye material with this biosynthetic indican, the scientists dissolved it in water and applied the solution alongside an enzyme that stripped away the glucose to re-form indoxyl. This indoxyl then spontaneously oxidized to leucoindigo. When removed from the liquid, the leucoindigo reacted with the air and turned to indigo.

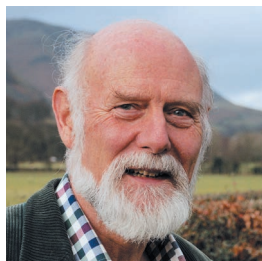
The clever mechanism goes further than previous attempts to clean the process, because it kills two polluting birds with one stone. First, it does away with the wasteful chemical synthesis of indigo.

Second, unlike previous indigo biosyntheses, this project removes the damaging bleaching stage that converts indigo to leucoindigo.

Industry churns out some 50,000 tonnes of synthetic indigo a year, and the bacterial system will need to be optimized and scaled up to make it commercially viable. The glucose molecules must be separated and removed, for one, and the enzyme used to liberate the indoxyl is expensive.

The scientists are optimistic that these challenges can be overcome. Are they right to be? One reason that biofuel production is cheap enough to be possible commercially is that it uses enzymes farmed from fungi. A useful step to prove the credentials of the greener denim dye would be to develop a similar low-cost way to make the required enzyme.

Still, indigo production has not always welcomed novelty. Until well into the eighteenth century, France protected its woad industry by threatening users of indigo imported from India and other foreign sources with the death penalty. But given that the popularity of blue denim shows no signs of slowing, the process that produces it sorely needs a new trend. ■



Funders should mandate open citations

All publishers must make bibliographic references free to access, analyse and reuse, argues David Shotton.

Over the past two decades, open access to journal articles, software and research data has changed from aspirational to commonplace. However, truly open scholarship also requires that bibliographic references be freely available for analysis and reuse.

Citations — the links created when a published work acknowledges other works in its bibliographic references — knit together independent works of scholarship into a global endeavour, and they are important for assigning credit to other researchers.

Analyses of citations can reveal how scientific knowledge develops over time and illuminate patterns of authorship. Such information is essential for assessing scholars' influence and making wise decisions about research investment. Bibliographic databases and citation indices are also crucial to individual researchers: they enable automated tools to hunt for relevant papers throughout the literature.

Making reference lists from articles free to view is insufficient for these purposes; to be useful, open references must be stored in a machine-readable format in a centralized repository. Crossref, the DOI-registration agency used by most academic publications, has provided such a repository since 2000, but its references are freely available only if publishers explicitly specify that they be made open. Funders and the scientific community must push harder for this.

Last year was eventful for open references. In April, more than 60 publishers (including Springer Nature) responded to a call from the Initiative for Open Citations (IIOC) — an effort that I co-founded — to unlock the reference lists of their scientific articles. By September, more than half of the nearly one billion journal-article references deposited at Crossref had been made open, up from 1% before IIOC launched. Bibliometric visualizations using this open data set have already appeared. They reveal, for instance, how co-authorship maps within particular disciplines and, at a larger scale, links between disciplines. In December, an open letter signed by more than 250 scientometricians called for publishers to open up their references (see go.nature.com/2crblo9). For reasons of both international equity and methodological integrity, scholars need access to comprehensive open reference data, and they need to be able to show the raw data behind their analyses.

That is presently not the case. The two most authoritative sources of citation data are Clarivate Analytics' Web of Science, which grew from the Science Citation Index created by Eugene Garfield in 1964, and Elsevier's Scopus, launched in 2004. Neither is open or comprehensive. Most research universities pay tens of thousands of dollars annually to access one or both of them, whereas institutions and independent scholars that cannot afford such a cost have no access.

However, the idea that references are proprietary data is fading. In addition to the half-billion references already made open by Crossref,

the OpenCitations Corpus, the repository I run with computer scientist Silvio Peroni, has already published 12.8 million citation links from PubMed Central under a Creative Commons waiver that puts them in the public domain. These are fully curated and semantically enhanced in Linked Open Data format to assist automated analysis.

Two significant barriers prevent comprehensive reference availability through Crossref. First, although it is easy to do so, two-thirds of Crossref's publisher-members, in particular the smaller ones, do not submit references along with the other details of their publications.

The second obstacle is created by publishers that submit references to Crossref, but do not make them open. Elsevier is by far the largest member of this group, which also includes the American Chemical Society, IEEE and Wolters Kluwer Health. Elsevier deposits about one-third of all journal-article references stored by Crossref, these constitute nearly

two-thirds of those that are not presently open.

The rationale for Elsevier not opening up its references is financial: free availability of its numerous bibliographic references would undermine Elsevier's ability to sell access to such data.

Companies such as Elsevier have invested considerable resources over many years into creating databases that can be used for bibliometric analyses. Elsevier argues that it is reasonable to charge for high-quality citation analysis, that curating citation data entails costs, including licensing fees, and that it cannot make reference lists from its journals freely available because it could not then afford to add value to these data.

However, I believe that Elsevier's decision not to open up its raw reference data is misguided.

Because it is bad for scholarship, it cannot be good in the long term for a business that seeks to serve scholars. In an increasingly open world, Elsevier's reputation will suffer, and its publications will become less visible. Instead, Elsevier executives should have more confidence in the advantage their analytical services give them in the citations market.

I call on all parties who could potentially benefit — including researchers, librarians, bibliometricians, funders, academic and research administrators, governmental agencies, members of the general public, and other stakeholders committed to open scholarship — to campaign for comprehensive open access to bibliographic references, and to actively develop, support and use services providing such access. However, where polite encouragement falls on deaf ears, sterner measures are required. Specifically, major funders should extend their open-access mandates and require grant recipients to publish only in journals whose publishers ensure their references are open. ■

David Shotton is co-director of OpenCitations, and a senior researcher at the Oxford e-Research Centre, University of Oxford, UK. e-mail: david.shotton@opencitations.net

SEVEN DAYS

The news in brief

RESEARCH

Views of science

Most people in the United States are in favour of science, but few are knowledgeable about how research is conducted, according to a survey by the advocacy group Research!America, based in Arlington, Virginia. The survey, released on 2 January, asked 1,005 people about their views of science and scientists. Although 82% of respondents thought that scientists were trustworthy, 81% could not name a living scientist and 67% could not name a research institution. About half of the respondents said they believed that great science will continue under US President Donald Trump's administration, and 67% agreed that public policies should be rooted in the best available science. Research!America's surveys have found similar results over the past decade.

HEALTH

A costly treatment

A gene therapy to treat hereditary blindness will cost US\$425,000 per eye, pharmaceutical company Spark Therapeutics announced on 3 January. The US Food and Drug Administration approved the treatment, called Luxturna (voretigene neparvovec), in December; it was the first US approval for a gene therapy that targets disease-causing mutations. Observers were keen to see what Spark, based in Philadelphia, Pennsylvania, would charge for the treatment, which is administered only once in each eye and could set a precedent for future gene therapies. At \$850,000 for a full, two-eye treatment, the cost is below the predicted price tag of \$1 million — but

has still raised eyebrows, given widespread concern over high drug prices.

EVENTS

Congress postponed

Organizers of the 105th Indian Science Congress have delayed the country's largest gathering of scientists until March amid concerns over the venue. At an emergency meeting on 27 December, the Indian Science Congress Association in Kolkata postponed the conference, which was scheduled for 3–7 January at Osmania University in Hyderabad. The association said in a statement that the event was postponed because

the university was no longer in a position to host the event "due to certain issues [on] the campus". No further details were given. The congress will now be held at Manipur University in northeast India on 16–20 March.

Dementia pull-out

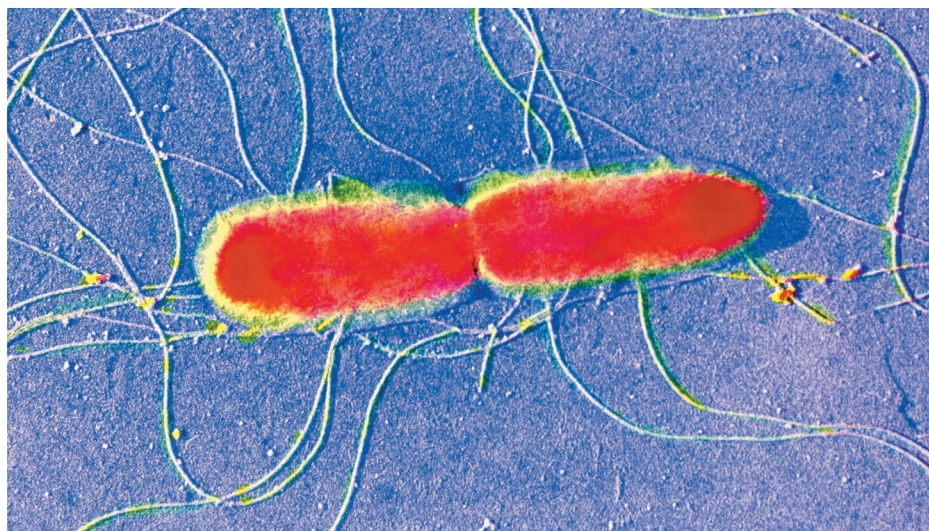
The multinational drug firm Pfizer will abandon research on dementia treatments, joining a stream of major pharmaceutical companies that have fled the high-risk research field in the past decade. On 6 January, the company said that it expects to shed 300 US jobs from its neuroscience discovery and early stage drug-development programmes

in Andover and Cambridge, both in Massachusetts, and in Groton, Connecticut. In 2012, Pfizer stopped a clinical trial of an antibody therapy for Alzheimer's disease because it demonstrated no clinical benefit. No therapies for Alzheimer's are yet available.

FACILITIES

AI industry hub

China will invest 13.8 billion yuan (US\$2.1 billion) in an industrial park devoted to artificial intelligence (AI). The park in western Beijing is expected to host 400 companies and a national laboratory to house collaborations between



CNRI/SPL

Improved typhoid jab gets go-ahead

A new vaccine against typhoid fever will be rolled out to millions of children in low-income countries, after the World Health Organization (WHO) announced its endorsement on 3 January. The product — developed by Bharat Biotech in Hyderabad, India — is a typhoid conjugate vaccine, which means that it provides longer-lasting protection and requires fewer doses than do other typhoid immunizations. The WHO's endorsement allows the vaccine to be

procured by United Nations agencies. Gavi, an organization based in Geneva, Switzerland, that funds vaccines for low-income countries, said it would spend US\$85 million on deploying the vaccine, with child immunizations likely to begin in 2019. Typhoid bacteria (*Salmonella Typhi*; pictured), which spread through contaminated food and water, cause an estimated 11 million to 20 million infections and 128,000–161,000 deaths each year.

NASA industry and domestic and foreign universities and research institutions. Last July, the central Chinese government released plans for the country to become the world leader in AI by 2030. In November, it announced that information-technology giants Baidu, Alibaba and Tencent would be partners in a national AI strategy. Google has also set up an AI research centre in Beijing.

ENERGY

Offshore drilling

The US Department of the Interior has reversed course on offshore drilling, proposing to open up most coastal waters for oil and gas development. The draft leasing programme for 2019–24, released on 4 January, would overturn extensive drilling restrictions put in place under former US president Barack Obama and allow for energy development on more than 90% of the US outer continental shelf. Under the proposed plan, the interior department would auction off 47 oil and gas leases over five years, including 16 leases along the east and west coasts — areas that have been off limits for federal leasing for more than three decades. The proposal opened to public comment on 8 January and faces opposition from environmentalists and many coastal states.

PEOPLE

UK science minister

Sam Gyimah was appointed UK minister for universities and science on 9 January, as part of a reshuffle of the Cabinet, the government's most senior decision-making body. Gyimah, who became a Member of Parliament in 2010, moves from the Ministry of Justice, where he was a junior minister. He campaigned for Britain to remain in the European Union in the 2016 referendum. The role of science minister will remain split between the Department for Education and the Department for Business, Energy and Industrial Strategy. Gyimah replaces Jo Johnson, who moves sideways to become a junior minister at the Department for Transport.

Astronaut dies

John Young, one of NASA's most experienced astronauts, died on 5 January from complications of pneumonia. He was 87. Trained as a test pilot, Young (pictured) first flew in space in 1965 aboard *Gemini 3*. He orbited the Moon in 1969 on *Apollo 10*, and landed on the lunar near-side in the Descartes highlands in 1972 with *Apollo 16*. In 1981, he commanded the space shuttle *Columbia* on its maiden flight; two years later, he was commander of the first



Spacelab mission, focusing on scientific experiments. Young was the first person to have flown into space six times.

French-agency head

France's Prime Minister Edouard Philippe has proposed Antoine Petit as the next president of the nation's main basic-science funder, the CNRS. Petit is currently chief of Inria, the country's research agency for computer science and applied mathematics. French President Emmanuel Macron endorsed the nomination on 3 January, but Petit will still need to be interviewed and approved by both houses of parliament before being formally appointed. With an annual budget of €3.3 billion (US\$3.9 billion), the CNRS is Europe's largest

basic-research agency. Petit would succeed Alain Fuchs, who left the job last October — four months before the end of his four-year term.

CLIMATE CHANGE

Extreme weather

Hurricanes, wildfires and other natural disasters caused a record US\$306 billion in damages in the United States in 2017, the US National Oceanic and Atmospheric Administration said on 8 January. Sixteen events each caused at least \$1 billion worth of damage, with Hurricane Harvey — which hit Texas in August — topping the list at \$125 billion. Other notable events included hurricanes Irma and Maria, wildfires in California and two tornado outbreaks in the central and midwestern United States. The previous record for damages — \$215 billion, adjusted for inflation — was set in 2005, the year that Hurricane Katrina devastated Louisiana, Mississippi and other parts of the US Gulf Coast.

Climate panel back

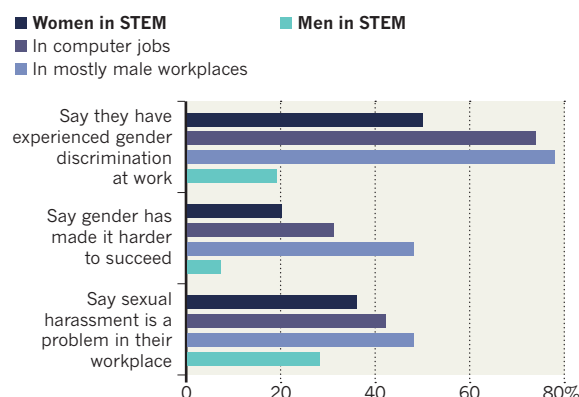
A federal climate advisory committee disbanded in August by US President Donald Trump is being revived. Columbia University and the state of New York are re-establishing the committee to help businesses and state and local governments make better use of the US National Climate Assessment, which is scheduled for completion this year. New York governor Andrew Cuomo, co-chair of an alliance of US states that is committed to action on climate change, announced the state's support on 2 January. Columbia University's Earth Institute in New York City is hosting the effort, and 10 of the original 15 committee members have agreed to serve, including climate scientist and former co-chair Richard Moss.

TREND WATCH

Half of women in science, technology, engineering and mathematics (STEM) jobs in the United States say that they have been discriminated against at work, according to a survey of nearly 5,000 people published on 9 January. That compares with 41% of women in other sectors. The survey also finds that more black people in STEM jobs (62%) than in other jobs (50%) report experiencing racial discrimination. Both groups say that negative stereotypes affect recruitment and promotion.

DISCRIMINATION IN STEM JOBS

In the United States, 50% of women in science, technology, engineering and mathematics (STEM) jobs report experiencing discrimination at work.



NEWS IN FOCUS

COMPUTING Groups race to develop silicon as a platform for quantum computers **p.136**

PALAEONTOLOGY Scratches on pterosaur teeth provide clues about the animals' diet **p.138**

CLIMATE SCIENCE Carbon storage in wetlands shows surprising consistency **p.139**

NEUROSCIENCE The tools to observe and manipulate memory get an upgrade **p.146**



WORLD VIEW



A Stratollite balloon made by World View is inflated in Page, Arizona.

SPACE SCIENCE

Scientific ballooning aims for the stratosphere

Commercial providers open the market for new types of research flight.

BY ALEXANDRA WITZE
BROOMFIELD, COLORADO

Private companies want to take scientific experiments sky-high in 2018 — aboard specialized balloons.

For decades, agencies including NASA and France's National Centre for Space Studies have flown balloon-borne experiments to realms higher than aeroplanes can reach but lower than satellites' orbits. Now, companies such

as World View of Tucson, Arizona, are lofting payloads quickly and cheaply into the stratosphere, between 16 and 30 kilometres up. The commercial balloon flights have new capabilities that open up fresh types of science — such as low-cost monitoring of natural disasters, or testing how to explore Venus by studying Earth's geology, says Alan Stern, a planetary scientist at the Southwest Research Institute in Boulder, Colorado, and a co-founder of World View.

"We're turning what was rare scientific ballooning into something routine," Stern says.

Balloons occupy a sweet spot between planes, which can survey small areas of land in great detail, and satellites, which span the globe but provide images at much lower resolutions. "We need observations from balloons, because they're just so powerful," said Karl Hibbits, a planetary scientist at the Johns Hopkins University Applied Physics Laboratory in Laurel, Maryland. He spoke at a meeting of the ▶

► Next-Generation Suborbital Research Conference in Broomfield, Colorado, on 18–20 December.

Among the ballooning companies that accept scientific payloads are Raven Aerostar of Sioux Falls, South Dakota, and Near Space Corporation of Tillamook, Oregon. World View has made a splash in the past year by developing a standardized ‘Stratollite’ platform that dangles beneath its balloons. A payload that might cost more than US\$1 million to fly on a NASA balloon could fly for tens of thousands of dollars on World View if it shares a Stratollite with other experiments, Stern says.

In 2018, World View aims to fly up to four times a month, says Jane Poynter, the company’s chief executive. Each balloon would lift one Stratollite carrying one or more experiments into the stratosphere. The longest flight so far lasted just over five days, but Poynter says the company hopes to conduct weeks-long flights in the near future.

Robert Grimm, a planetary scientist at the Southwest Research Institute, flew an experiment on a World View balloon in October to test designs for a possible mission to Venus. The planet’s surface is too hot for equipment to survive for long, but conditions high in Venus’s atmosphere are much more temperate — meaning that scientists could use balloons as a way to study the planet for months, rather than minutes or hours.

After taking off in Idaho, the balloon soared for 500 kilometres before touching down in Montana. Like a high-flying metal detector, Grimm’s on-board equipment measured changes in electrical properties within a granite-rich mountain range below. Collecting such data over Venus could illuminate the geology at or beneath the planet’s surface, says Grimm, who hopes to fly further experiments in May.

World View has also developed ways to hold its balloons nearly stationary over a point of interest. The company directs the balloon up and down to catch the wind and keep the craft in approximately the same location. Google’s parent, Alphabet, uses a similar approach to keep its Project Loon balloons in one spot. The company has been testing whether the balloons, built by Raven Aerostar, can provide Internet connectivity in places such as Puerto Rico, following last September’s devastating Hurricane Maria.

NASA is developing advanced balloon technology for scientists, including its ‘super-pressure’ balloons that can fly for up to 100 days — a period suitable for long-term studies such as certain astronomical observations. But the work is expensive and technologically challenging. For many experiments, “the World

View flights are actually there”, says Thomas Zurbuchen, NASA’s associate administrator for science in Washington DC. “We’re really interested in doing some science on these new platforms.” (The agency funded Grimm’s World View flight.)

To Adrienne Dove, a planetary scientist at the University of Central Florida in Orlando, stratospheric balloons offer a new opportunity to explore the physics behind spaceflight. She studies how dust clumps together in low-gravity conditions — important for lunar and planetary exploration — and has worked with sounding rockets and the ‘vomit comet’ aeroplanes that create low gravity for short periods during their parabolic flights. “My interest is in developing microgravity capability on balloons, which currently doesn’t exist,” she says.

Looking even further into the future is Siddharth Krishnamoorthy, an aerospace engineer at the Jet Propulsion Laboratory in Pasadena, California. His team wants to use stratospheric balloons to listen for low-frequency infrasound signals coming from earthquakes, as a test for possible future missions to probe for seismic activity on Venus.

That would mean floating in the stratosphere above earthquake-prone places such as Oklahoma or California, listening for infrasound signals and pretending they are on Venus. “Yes, it’s cool,” says Krishnamoorthy. ■

“We need observations from balloons, because they’re just so powerful.”

COMPUTING

Silicon gains ground in quantum-computing race

Slow-starter seeks to catch up with rival techniques.

BY DAVIDE CASTELVECCHI

In the next few weeks, a research group at the Delft University of Technology in the Netherlands expects to receive an important package. Its contents promise to increase competition in the race to produce useful quantum computers.

Shipped from the research-and-development facilities of semiconductor giant Intel in Hillsboro, Oregon, the parcel holds the first quantum computer manufactured with the techniques used to fabricate silicon chips in conventional computers. Although the silicon method currently lags behind other approaches to building quantum computers, the company hopes that the technique could accelerate the development of devices that go beyond proof-of-concept

curiosities, says James Clarke, who heads Intel’s quantum-hardware development. “I think you’ll hear a lot about silicon quantum computing this year,” Clarke says.

The relatively modest device represents the latest move in the push to give silicon a boost over other approaches. Some scientists also see promise in the silicon route. Physicists such as Michelle Simmons at the University of New South Wales (UNSW) in Sydney, Australia, are developing their own ways of building quantum computers using silicon. In May 2017, she founded an Aus\$83 million (US\$65 million) start-up called Silicon Quantum Computing, backed in part by the Australian government.

Quantum computers aim to exploit two small-scale phenomena to outperform their classical counterparts, which encode bits of

information as 0s and 1s. In the quantum world, units of information are called qubits, and each qubit can exist simultaneously in a ‘superposition’ of both 0 and 1. Two bits can also be entangled, so that the state of one qubit determines the state of its partner. This enables quantum devices to conduct calculations in parallel.

Physicists in many labs have developed prototype quantum computers, which often operate at temperatures close to absolute zero. The frontrunners in the race use one of two methods to encode the qubits: single ions held in traps, or oscillating currents in superconducting loops. Both systems require exquisite control: the ion technique uses complex laser systems to read and write each qubit, and superconducting qubits must each have a device to control them using radio waves.

PUBLISHING

Elsevier grants a reprieve

It allows German institutions continued journal access.

BY QUIRIN SCHIERMEIER

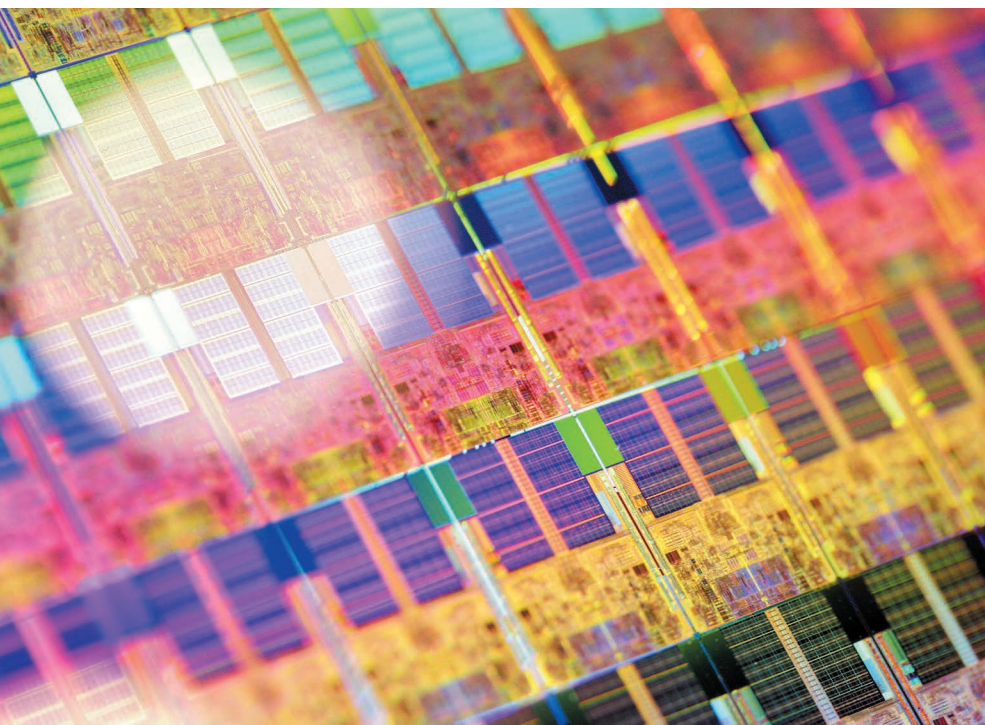
The Dutch publishing giant Elsevier has granted uninterrupted access to its paywalled journals for researchers at around 200 German universities and research institutes that had refused to renew their individual subscriptions at the end of 2017.

The institutions had formed a consortium to negotiate a nationwide licence with the publisher. They sought a collective deal that would give most scientists in Germany full online access to about 2,500 journals at roughly half the price that individual libraries have paid in the past. But talks broke down and, by the end of 2017, no deal had been agreed. Elsevier now says that it will allow the country's scientists to access its paywalled journals without a contract until either a national agreement is reached or 200 individual contracts are hammered out.

The two sides had “constructive conversations well into December”, says Harald Boersma, a spokesman for Elsevier. “We will continue our conversations in the first quarter of 2018 to find an access solution for German researchers in 2018 and a longer-term national agreement,” he says. “Where access agreements ended, we have informed these institutions that we would maintain access to our content while we continue to work with the German Rectors' Conference [which leads negotiations for the consortium] on a solution and specifically a one-year extension to existing contracts, covering 2018.”

Günter Ziegler, a mathematician at the Free University of Berlin and a member of the consortium's negotiating team, says that German researchers have the upper hand in the talks. “Most papers are now freely available somewhere on the Internet, or else you might choose to work with preprint versions,” he says. “Clearly our negotiating position is strong.”

Academic-publishing experts around the world are keenly observing the situation in Germany. The nationwide deal sought by scientists includes a open-access option, under which all corresponding authors affiliated with German institutions would be allowed to make their papers free to read and share for anyone in the world. This would be a milestone for global efforts to make the results of publicly funded research immediately and freely available to scientists and the wider public, they say. ■



Techniques for fabricating conventional silicon chips could be used to make quantum devices.

YOSHIKAZU TSUNO/AFP/GETTY

Proponents of the silicon technique see major advantages in using a semiconductor to code qubits. They can be manipulated much more simply using microscopic electric leads etched right onto the chip. And if the same large-scale manufacturing techniques for making chips could be transferred to the quantum realm, it could become easier to turn the technology into commercial products.

A LONG ROAD

The idea of building quantum computers out of silicon is not new. Bruce Kane, an experimental physicist now at the University of Maryland in College Park, first suggested encoding qubits in the magnetic orientation, or ‘spin’, of phosphorus nuclei embedded in silicon 20 years ago¹. At about the same time, David DiVincenzo, a theoretical physicist then at IBM in Yorktown Heights, New York, and his collaborator Daniel Loss at the University of Basel in Switzerland proposed a way of storing information in the spins of mobile electrons inside semiconductors². Both proposals led to a number of experimental demonstrations but, for a long time, the quality of the materials limited progress.

Building a quantum computer using silicon took years of “not very flashy” developments in materials science and engineering, says physicist Jason Petta of Princeton University in New Jersey. Physicists at the UNSW Centre for Quantum Computation and Communication Technology, which Simmons directs, have done much of that groundwork. And Simmons developed a manufacturing technique that requires fewer control leads, preventing inevitable issues of crowding once quantum

devices scale up, she says. “I want to engineer everything out that isn't essential and make things as simple as possible.”

In 2017, two groups reached a milestone when they designed the first fully controllable two-qubit devices in silicon. Petta and his collaborators achieved that feat³, as did a separate team⁴ led by Lieven Vandersypen at Delft.

Intel, which is investing US\$50 million over 10 years at Delft, is now manufacturing multiple-qubit electron-spin devices for Vandersypen, in the same type of factory where it develops microprocessor-fabrication techniques. Industrial partners can help by providing reliably identical devices, he says.

“We hope that we can accelerate spin qubits to compete” with the more mature approaches, Clarke says. Simmons' start-up aims to build a ten-qubit machine within five years. Google, IBM and a number of other companies and academic labs are all using different techniques to build quantum computers with around 50 superconducting qubits — and so is Intel itself, which is hedging its bets by supporting more than one technical approach. ■

1. Kane, B. E. *Nature* **393**, 133–137 (1998).
2. Loss, D. & DiVincenzo, D. P. *Phys. Rev. A* **57**, 120–126 (1998).
3. Zajac, D. M. *et al.* *Science* <http://dx.doi.org/10.1126/science.aao5965> (2017).
4. Watson, T. F. *et al.* Preprint at <https://arxiv.org/abs/1708.04214> (2017).



Fossil teeth of *Dimorphodon macronyx* suggest that it ate insects and small land vertebrates.

PALAEOLOGY

Pterosaur teeth reveal ancient diet

Feeding habits of flying reptiles have been much debated.

BY JOHN PICKRELL

Microscopic scratches on fossil teeth are forcing palaeontologists to rethink some cherished ideas about the diets of pterosaurs — flying reptiles that ruled the skies while terrestrial dinosaurs flourished on the lands beneath them.

Since pterosaur fossils were first uncovered in the eighteenth century, researchers have made assumptions about their eating habits, mostly from indirect clues such as the shapes of their teeth and the environments they lived in. But Jordan Bestwick, a palaeontologist at the University of Leicester, UK, and his colleagues sought more-direct evidence: they performed the first examination of fossilized

pterosaur teeth for tiny abrasions caused by food. Microscopic scratches and chips create characteristic surface textures that vary according to an animal's diet, says Bestwick.

The preliminary findings offer new details about the feeding habits of some species, and confirm theories about the diets of others. Bestwick presented the results, which will form part of his PhD thesis, at the Palaeontological Association's annual meeting in London on 18 December.

One surprise finding in the analysis raised questions about the pterosaur *Dimorphodon macronyx*, which researchers assumed had hunted fish. The wear and tear on the reptile's teeth suggests that it actually feasted on insects and land vertebrates.

Although pterosaurs existed for 150 million years, complete fossils are relatively rare, and gut contents have been recovered from just four species. That means that most hypotheses about species' diets have been "little more than speculation based on scant evidence", says Bestwick.

He and his colleagues have so far examined 11 pterosaur species, looking at tooth specimens held at institutions such as the Natural History Museum in London and the Museum für Naturgeschichte in Berlin. They used infinite-focus microscopes to create 3D images of tooth wear. They then used statistical methods to look at wear patterns on pterosaur teeth, alongside the teeth of living species of bats, lizards and crocodilians that are known to eat insects or fish and other vertebrates.

Analysis of the pterosaur *Rhamphorhynchus* reveals wear patterns that are statistically similar to those seen in modern relatives of crocodiles. This suggests that *Rhamphorhynchus* ate fish, backing up a long-standing hypothesis about the pterosaur's diet, Bestwick says. Wear patterns on the teeth of *Pterodactylus*, the first pterosaur ever described, in 1784, suggests that it was an omnivore, as some experts had also hypothesized, he adds.

Stephen Brusatte, a palaeontologist at the University of Edinburgh, UK, says the study is one of the first attempts to use a rigorous statistical method to determine what these flying reptiles ate. "This is a great example of how a combination of cutting-edge techniques and careful comparisons to modern species can help us understand how long-extinct animals behaved," says Brusatte.

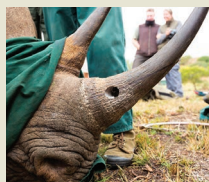
Steven Vidovic, a vertebrate palaeontologist at the University of Portsmouth, UK, says that complete fossils of pterosaurs are so rare because their light, hollow bones were relatively fragile and unlikely to fossilize. The lack of direct evidence of their diets has often led to researchers making assumptions on the basis of the reptiles' environment, he says. For instance, pterosaur remains are often found in coastal environments, which led researchers to assume that many species ate fish, he says.

Vidovic says the latest analysis will enable palaeontologists to test theories about pterosaurs' diets. "This new method presents a real opportunity to observe the hardness and abrasiveness of the food pterosaurs were consuming, and test hypotheses of ecology," he says. ■

MARK P. WITTON/CC BY 4.0


**MORE
ONLINE**

TOP NEWS



Rhinoceros poachers prosecuted using DNA database go.nature.com/2akgkny

MORE NEWS

- Long-awaited US report charts course for studies of Earth from space go.nature.com/2cht2rg
- How the immune system could stymie some CRISPR gene therapies go.nature.com/2d9zezy

NATURE PODCAST



Tabletop physics; what a memory looks like; and conflict's toll on wildlife nature.com/nature/podcast

MARTIN HARVEY/GALLO/GETTY

PUBLISHING

Indonesian preprint server takes off

Website's creators aim to open up the country's science to a wider audience.

BY IVY SHIH

A preprint server that focuses exclusively on Indonesian research passed a milestone on 5 December when the number of papers posted on it reached 1,500. INA-Rxiv is one of the first preprint repositories to focus on the work of a single country.

"I didn't think it would be this huge in such a short period of time," says hydrogeologist Dasapta Irawan, who helped to create INA-Rxiv, which launched in August.

Most preprint servers specialize in particular academic disciplines — including the original arXiv, which covers physics and mathematics. The four researchers who developed INA-Rxiv built it to draw attention to Indonesian research, which they felt was going unnoticed by the international science community. "I want people to understand that in Indonesia, we can produce original research and papers," says Irawan, who is based at the

Bandung Institute of Technology in Indonesia.

The server hosts papers in multiple disciplines — most in the natural sciences, followed by engineering, the social and behavioural sciences and arts and humanities — and accepts material written in Bahasa Indonesian and English. It operates in partnership with the Open Science Framework, a service run by the Center for Open Science in Charlottesville, Virginia.

Computer scientist Robbi Rahim at Indonesia's Medan Institute of Technology has uploaded 26 papers. One of those articles, about multimedia learning in mathematics and written in Bahasa, has been downloaded some 330 times. Rahim says that the site helps him to reach a big audience, because he can upload articles in both languages.

Irawan says that some Indonesian scientists

seem to be using INA-Rxiv to boost the chance of having their papers included in the government's new research-evaluation system, called the Science and Technology Index (SINTA). Launched in January 2017, SINTA ranks researchers and institutions by various metrics, including the number of publications listed in major citation databases and Google Scholar.

But Irawan says that SINTA does not index many open-access Bahasa-language journals. Some researchers, he says, seem to use INA-Rxiv to get around SINTA's limitation. That's because articles on the preprint server are automatically indexed on Google Scholar.

Although Indonesian scientists have embraced INA-Rxiv, some question whether it will improve the country's research. Psychology researcher Dicky Pelupessy of the University of Indonesia in Depok says that research quality is one of the reasons Indonesian scientists struggle to get their research read and cited internationally. ■

"In Indonesia, we can produce original research and papers."

CLIMATE SCIENCE

'Blue carbon' defies expectations

Results of soil survey could bolster efforts to monitor and protect wetlands around the globe.

BY JEFF TOLLEFSON

Tidal wetlands come in many forms, but they could be more alike below the surface than anyone realized. Whether it's a mangrove forest in Florida, a freshwater swamp in Virginia or a saltwater marsh in Oregon, the amount of carbon locked in a soil sample from each of these coastal ecosystems is roughly the same.

That's the surprising message from a new analysis of some 1,900 soil cores collected around the United States during the past few decades. "In terms of carbon stocks, all tidal wetlands are very, very similar," says Lisamarie Windham-Myers, an ecologist with the US Geological Survey (USGS) in Menlo Park, California, who is leading a 3-year, US\$1.5-million assessment of coastal

carbon funded by NASA. "The variability that everybody expected just doesn't exist."

Her team presented its findings last month in New Orleans, Louisiana, at a meeting of the American Geophysical Union; the researchers plan to publish data from 1,500 soil cores online as early as this month, and hope to release information on the remaining 400 later this year.

The discovery could bolster efforts to assess and protect the world's coastal wetlands. These ecosystems accumulate vast stocks of carbon that escape into the atmosphere when wetlands are destroyed. Development alters some 800,000 hectares of coastal wetlands around the world each year, sending roughly 500 million tonnes of carbon dioxide into the atmosphere — double the carbon emissions of Spain in 2016.

Over the past decade, scientists and policymakers have pushed to protect the carbon stored in coastal wetlands, known as blue carbon. The goal is to address climate change while protecting ecosystems that sustain fisheries, improve water quality and protect coastlines against storms. But raising money to support such efforts often requires determining precisely how much carbon these ecosystems hold, and how it accumulates over time.

Windham-Myers's team reanalysed raw data from some 1,500 sediment cores collected over the past several decades, and 400 newer samples. The data showed a clear relationship: the density of soils decreased as the fraction of carbon in those soils increased, and vice versa. As a result, the amount of carbon in any given cubic metre of soil remained roughly the same, regardless of differences in vegetation, climate, topography or water chemistry across blue-carbon ecosystems.

"It's almost like a universal constant," says Stephen Crooks, an independent geomorphologist in San Francisco, California, who analysed blue-carbon stocks in the latest US inventory of greenhouse-gas emissions and sinks. That report, which the US Environmental Protection Agency released in April last year, found that the United States' 3.8 million hectares of coastal wetlands soak up 8.1 million tonnes of CO₂ each year.

Estimates from a century's worth of soil surveys by the US Department of



Tidal wetlands such as this marsh in Oregon can store large amounts of carbon.

► Agriculture (USDA) showed more variation, but those figures were based on data collected by people who were often thinking more about agriculture on land. In the Mississippi delta, for instance, many early measurements were limited to surface sediments that are rich in carbon, and estimates of the soil density below the surface may have been too high. As a result, Windham-Myers says, the USDA

overestimated carbon stocks in the region.

Crooks says that if soil measurements from wetlands elsewhere agree with the US findings, global estimates of carbon stocks could improve. Windham-Myers and her colleagues recently examined data from coastal wetlands across Africa, and the results were consistent with the team's analysis of cores from US tidal wetlands.

But understanding how much carbon is in the ground is just a prelude to determining the rate at which wetlands sequester carbon. That figure depends in part on local topography and on the rate at which seas rise and create more space for carbon-rich sediments to accumulate. And methane emissions vary widely depending on whether water in a wetland is salty, fresh or brackish. Similarly, understanding how much carbon enters the atmosphere when a wetland is drained for agriculture or other purposes requires a more detailed understanding of the soil make-up. All of this information must be plugged into models to project how wetlands will evolve in the coming decades.

Crooks hopes that providing better data on the carbon stored by wetlands will encourage governments to halt the destruction of these ecosystems. "It's important that we find every mechanism that we can to offset our carbon emissions," Crooks says. "This is one piece of the puzzle." ■

CORRECTION

The image of the eclipse in 'Images of the year' (*Nature* **552**, 308–313; 2017) was an artistic representation that did not accurately depict the event. It has been replaced with a new image online.



TABLETOP PHYSICS PUSHED TO THE EDGE

BY GABRIEL POPKIN

It's possible that no one knows the electron as well as physicist Gerald Gabrielse. He once held one in a trap for ten months to measure the size of its internal magnet. When it disappeared, he searched for two days before accepting that it was gone. "You get kind of fond of your particles after a while," he says.

And Gabrielse has had ample time to become fond of the electron. For more than 30 years, he has been putting sophisticated electromagnetic traps and lasers to work to reveal the particle's secrets, hoping to find the first hints of what's beyond the standard model of particle physics — the field's long-standing, but incomplete, foundational theory. Yet for many of those years, it seemed as if he was working

in the shadow of high-energy facilities such as the Large Hadron Collider (LHC), the 27-kilometre-circumference, US\$5-billion particle accelerator near Geneva, Switzerland. "There was a time in my career when there weren't very many people doing this kind of thing, and I wondered if it was the right choice," he says.

Now, he's suddenly moving from the fringes of physics to the limelight. Northwestern University in Evanston, Illinois, is about to open a first-of-its-kind research institute dedicated to just his sort of small-scale particle physics, and Gabrielse will be its founding director.

The move signals a shift in the search for new physics. Researchers have dreamed of finding subatomic particles that could help

Researchers adapt atomic-physics tricks to look for evidence of new particles.

them to solve some of the thorniest remaining problems in physics. But six years' worth of LHC data have failed to produce a definitive detection of anything unexpected.

More physicists are moving in Gabrielse's direction, with modest set-ups that can fit in standard university laboratories. Instead of brute-force methods such as smashing particles, these low-energy experimentalists use precision techniques to look for extraordinarily subtle deviations in some of nature's most fundamental parameters. The slightest discrepancy could point the way to the field's future.

Even researchers long associated with high-energy physics are starting to look to low-energy experiments for glimpses beyond the

ALYSSA SCHUKAR FOR NATURE

Gerald Gabrielse in his low-energy-physics lab at Northwestern University in Evanston, Illinois, with postdoc Wayne Huang.

standard model. If such hints emerge, they could point the way to explaining the mysteries of dark matter and dark energy, which collectively constitute some 95% of the Universe. “This is sort of a tectonic shift in the way we think of doing physics,” says Savvas Dimopoulos, a theorist at Stanford University in California.

SQUASHED SPHERE

In some ways, these small-scale experiments are a return to how particle physics was once done. Gabrielse drew particular inspiration from a 1956 experiment by physicist Chien-Shiung Wu. In a laboratory at what is now the US National Institute of Standards and Technology in Gaithersburg, Maryland, Wu found an asymmetrical spatial pattern in how radioactive cobalt-60 atoms emit electrons. The finding, along with theoretical work, confirmed that two particles discovered almost a decade before were actually one and the same. It also helped to solidify faith in the burgeoning theoretical framework for the Universe’s fundamental particles and most of its fundamental forces, which would soon evolve into the standard model.

But physics was already moving towards bigger and more-expensive experimental machinery. Buoyed by a flush of post-Second World War cash and prestige, and by predictions that new particles would emerge in high-energy collisions, physicists proposed increasingly powerful and expensive particle accelerators. And they got them: facilities sprung up at Stanford; at Fermilab near Batavia, Illinois; at CERN near Geneva; and elsewhere. Quarks, muons, neutrinos and, finally, the Higgs boson were discovered. The standard model was complete.

And yet, as a description of the Universe, it is incomplete. The standard model doesn’t explain, for example, why antimatter and matter were not created in equal parts at the start of the Universe. If they had been, they would have annihilated each other, leaving behind a featureless void. The standard model also says nothing about dark matter, which seems to bind galaxies together, or about the dark energy that is pushing the Universe apart at an accelerating rate. “I like to call the standard model the great triumph and the great frustration of modern physics,” says Gabrielse. On the one hand, he says, it lets physicists predict some quantities “to ridiculous accuracy. On the other hand, we have a hole we can drive the Universe through.”

Gabrielse’s work trapping and probing particles at very low energies has taken him to a smaller facility at CERN, home of the LHC, to hunt for differences between matter and antimatter (see *Nature* 548, 20–23; 2017). He and his colleagues have produced the most precise measurement yet of a physical quantity — the size of the electron’s internal magnet, or spin¹.

But one of his biggest focuses in the past

decade has been pinning down the shape of the electron. Although it is usually seen as a simple point with negative charge, the electron could have hidden complexity. If certain symmetries of nature — rules that say the Universe behaves the same under various reversals — are violated, the electron’s charge won’t have a perfectly spherical distribution. Instead, virtual particles that constantly wink in and out of existence will skew the overall distribution of charge, squashing it slightly out of shape and giving it what physicists call an electric dipole moment, or EDM (see ‘Searching the particle sea’).

The standard model predicts a tiny squashing — so small, Gabrielse says, that “there’s essentially no hope to measure it in my lifetime”. But

“THIS IS SORT OF A TECTONIC SHIFT IN THE WAY WE THINK OF DOING PHYSICS.”

some theories posit as-yet-undetected particles that could make the electron’s EDM roughly one billion times larger. Many of those theories fall into a class called supersymmetry, an extension of the standard model that could explain why the Higgs boson’s mass is smaller than expected, and that could unify the electromagnetic, weak and strong forces in the early Universe. It might also reveal the nature of dark matter.

Attempts to measure the electron’s EDM go back more than four decades. Physicists have taken advantage of the fact that an electron with an EDM can rotate, or precess, around an electric field, tracing out a loop. The stronger the electric field, the faster — and more easily detectable — the precession.

But complications abound. Experimentalists can’t work with solitary electrons, because a strong electric field would cause them to skitter away. Luckily, atoms and molecules effectively lock electrons in place — and can produce internal electric fields stronger than the strongest laboratory-made field. Because atoms and molecules absorb light at specific frequencies, researchers can use lasers to trap and cool them — and nudge their internal electrons into different configurations.

By the mid 2000s, several generations of experiments building on these techniques had ratcheted down the upper limit on the size of the electron’s EDM, but not quite to the level that would reveal the influence of particles predicted by supersymmetry or other extensions of the standard model. One of those

experiments was conducted at Yale University in New Haven, Connecticut, by physicist David DeMille and his colleagues, using thallium ions². But DeMille was running out of ideas for teasing more accuracy from his experiment, which was demanding an increasingly byzantine arrangement of highly calibrated lasers, vacuum chambers and cryogenics.

A breakthrough came in 2008, when two theorists at JILA, a research institute in Boulder, Colorado, reported³ that the molecule thorium oxide had an internal electric field roughly 1,000 times the strength of thallium’s, which would make a precession effect in its electrons much easier to see. Around the same time, Gabrielse — who was then at Harvard University in Cambridge, Massachusetts — had wrapped up a long-running study and decided that he wanted to get into the electric-dipole game. He talked to John Doyle, also a physicist at Harvard, who had invented a new way to make focused beams of cold, slow-moving molecules. DeMille also contacted Doyle, and the three decided to join forces. In 2009, the trio’s experiment, called Advanced Cold Molecule Electron EDM, or ACME, received a 5-year, \$6.2-million grant from the US National Science Foundation.

PRECESSION PRECESSION

The group set up shop at Harvard. Gabrielse worked on making the team’s lasers — eight in total — more stable and accurate. Doyle focused on producing high-quality beams of thousands of thorium oxide molecules. And DeMille designed a system to align the molecules and shield them from outside interference.

In the experiment, a lab-made electric field orients the thorium oxide molecules. A pair of lasers then sets the spin direction of an electron inside each molecule to be perpendicular to the molecule’s internal electric field, and a magnetic field is used to make the particle’s spin precess. If the electron has an EDM, it will slightly add to or subtract from that rotation. After about one millisecond, polarized laser light bouncing off the molecules reveals how far their electrons have precessed. The experiment is then repeated with the molecules’ orientations reversed, which should reverse the direction of precession due to an EDM. The larger the difference in precession angle, the larger the EDM.

In early 2014, the researchers reported⁴ that they had not seen evidence for an EDM in their set-up, which was sensitive to an angular difference of about 100-millionths of a degree. That drove the upper limit of the electron EDM down by more than a factor of 10, to 8.7×10^{-29} in units of centimetres multiplied by electron charge. If an electron were the size of Earth — and Earth a perfect sphere — the limit would correspond to moving a patch of material roughly 20 nanometres thick from one pole to the other.

The ACME team argued that the result has big implications for theories beyond the standard model, nixing many hypothetical supersymmetric particles that would exist in an

energy range probed by the LHC. But some theorists counter that plenty of remaining theories — supersymmetric and otherwise — predict an electron EDM smaller than those ruled out by the ACME team. Gabrielse finds the surviving theories more and more contrived. “Theorists are wily,” he says. “Every time we exclude something, they try to wiggle out.”

ACME is not alone in this effort. After earning a Nobel prize in 2001 for creating a new phase of matter called a Bose–Einstein condensate, JILA physicist Eric Cornell teamed up with Jun Ye, also at JILA, to look for an EDM. Rather than manipulate molecules as they pass by in a beam, as ACME does, Cornell and Ye decided to use a rotating electric field to trap molecular ions with large internal fields, giving electron precessions longer to reveal themselves. DeMille calls the idea “brilliant and far from obvious”.

Cornell faced a setback when he lost an arm to necrotizing fasciitis in 2004. But it led to a joke he likes to tell when he gives talks: “His left sleeve is empty, and he’ll say, ‘If anybody should know about asymmetry, it’s me,’” says former lab mate Chris Monroe, now a physicist at the University of Maryland in College Park. After a decade building and refining what Cornell calls a “two-tabletop experiment” (because it occupies two tables in his lab), he and his co-authors finally published their first results last year⁵, coming within a factor of 1.5 of ACME’s 2014 limit. “I might not have started if I had realized how hard it would be,” says Cornell.

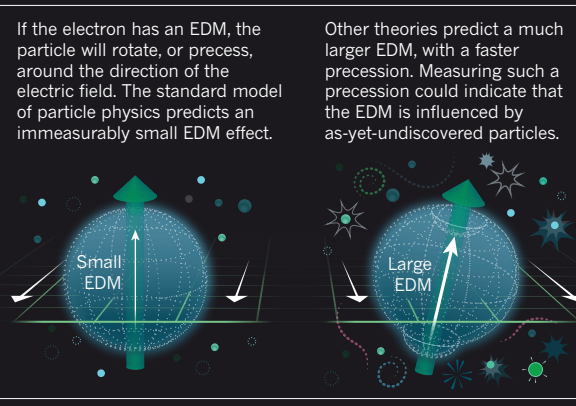
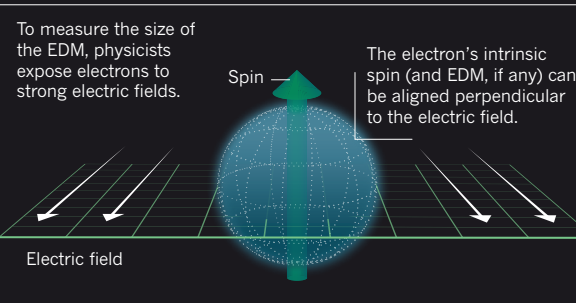
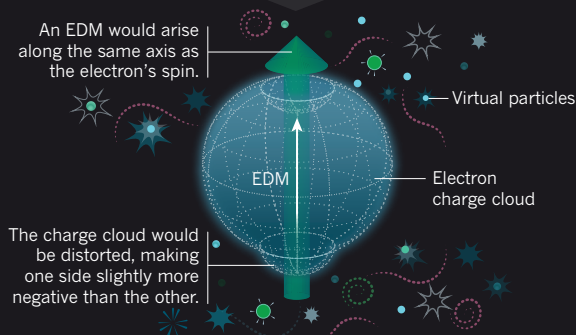
Now, researchers are closing in on new EDM results. The ACME physicists have increased the number of molecules they can send into their experimental apparatus by a factor of 400. They expect this and other improvements to sharpen the experiment’s precision by a factor of ten — allowing them to hunt for effects beyond the energy range of the LHC. The JILA team is also gearing up for experiments set to push beyond the LHC’s reach. And researchers at Imperial College London who held a former electron-EDM measurement record⁶ have plans for experiments with laser-cooled ytterbium monofluoride molecules; they hope their test will be 1,000 times more precise than ACME’s first run.

The electron isn’t the only low-energy peephole into the world beyond the standard model. Some physicists are searching for EDMs in neutrons or atoms, which, like the electron, could reveal a violation of one of nature’s symmetries. Others are adapting an entirely different

SEARCHING THE PARTICLE SEA

Physicists are hunting for evidence that the electron’s charge cloud might be not be perfectly round, which could indicate the presence of new particles.

The electron moves through a sea of virtual particles that are constantly popping into and out of existence. According to many theories, these should distort the electron’s charge cloud, creating a corresponding property called an electric dipole moment (EDM).



technology in service of fundamental physics: atomic clocks. The frequencies of radiation absorbed and emitted by the atoms that make up these clocks depend only on certain fundamental constants of nature. A slight deviation in those frequencies could lend support to theories that attempt to explain why gravity is so much weaker than the Universe’s other forces.

The ability to test this idea was out of reach until the early 2000s, when researchers developed atomic clocks that operate in the optical range of the electromagnetic spectrum instead of in the microwave. Their higher frequencies meant that time could be sampled at a much higher rate, enabling the creation of clocks so precise that they would lose or gain less than

one second over the age of the Universe. Researchers have since used data from such clocks to search for changes in the ratio between the electron’s and proton’s masses and in the fine-structure constant — a fundamental parameter that governs the strength of the electromagnetic force. Others, following a proposal⁷ by Asimina Arvanitaki, a theorist at the Perimeter Institute for Theoretical Physics in Waterloo, Canada, are using clocks to look for subtle oscillations that might be created by a hypothesized dark-matter candidate called the axion, or a related particle.

So far, these investigations have yielded no new physics. But they show how a younger generation of physicists is infusing the field with new ideas, says Dimopoulos, who was Arvanitaki’s PhD adviser. “There’s a lot of theoretical ideas that have been, in a sense, overlooked because everybody was focusing on the LHC and the previous colliders,” he says.

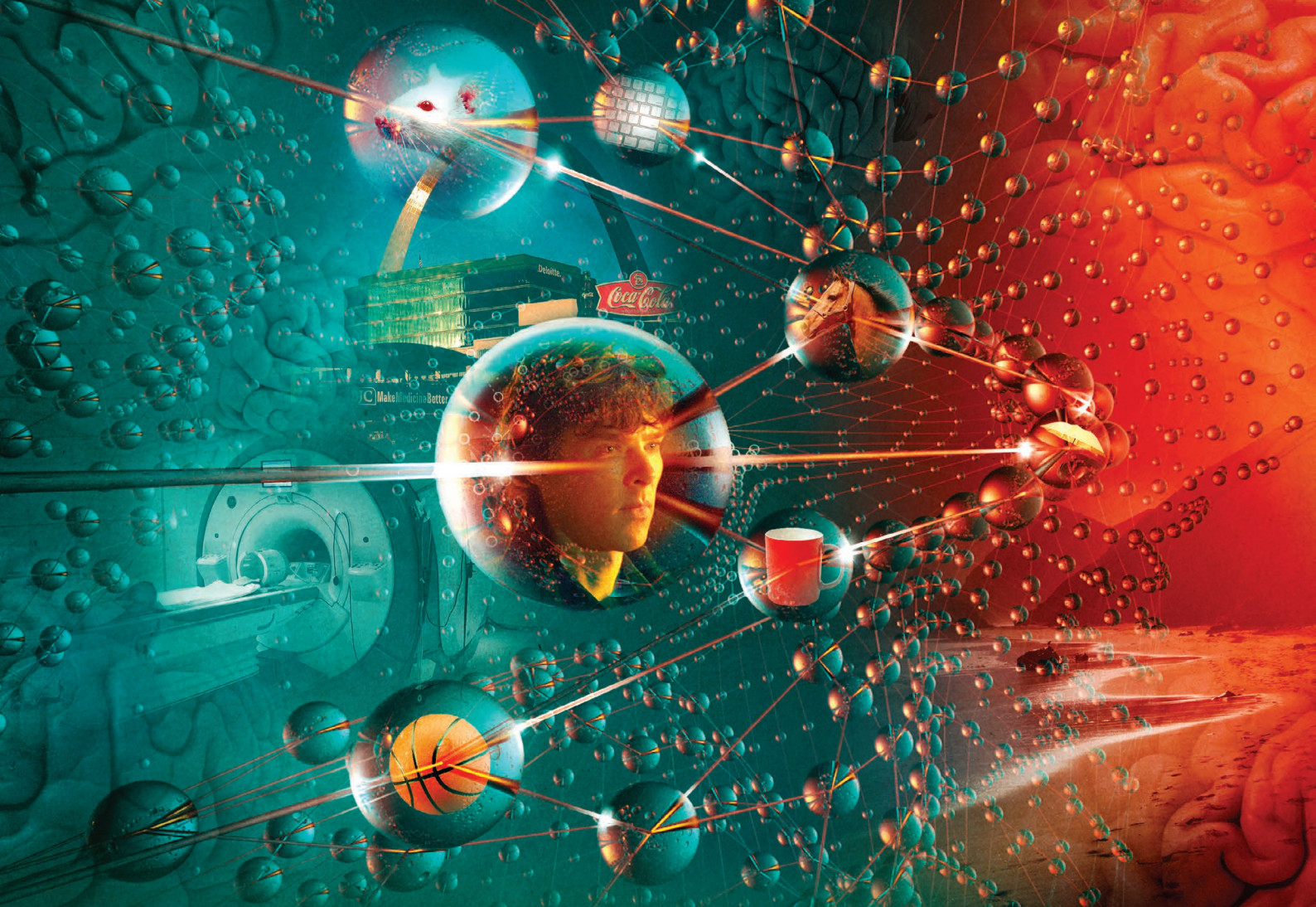
No one expects such tabletop experiments to replace particle colliders. Rather, they could guide physicists to the right energy range for more detailed study. Right now, the collider community suspects that it needs more energy than the LHC is designed to reach, but it’s unclear how much will be sufficient. Findings from low-energy experiments might influence a multibillion-dollar decision about the next big collider, and that has put added pressure on researchers working in this tabletop realm. “We have to do almost everything with more care than is typical in the standard atomic-physics experiment,” says DeMille.

Gabrielse has high hopes for the team’s next experiment — and for the work at his centre at Northwestern, which is set to open this year. But he can make no promises. “We’re fishing for a fish whose shape and colour and

speed and equipment for biting are completely unknown.” ■

Gabriel Popkin is a freelance journalist based in Mount Rainier, Maryland.

1. Hanneke, D., Fogwell, S. & Gabrielse, G. *Phys. Rev. Lett.* **100**, 120801 (2008).
2. Regan, B. C., Commins, E. D., Schmidt, C. J. & DeMille, D. *Phys. Rev. Lett.* **88**, 071805 (2002).
3. Meyer, E. R. & Bohn, J. L. *Phys. Rev. A* **78**, 010502(R) (2008).
4. The ACME Collaboration. *Science* **343**, 269–272 (2014).
5. Cairncross, W. B. *et al. Phys. Rev. Lett.* **119**, 153001 (2017).
6. Hudson, J. J. *et al. Nature* **473**, 493–496 (2011).
7. Arvanitaki, A., Huang, J. & Van Tilburg, K. *Phys. Rev. D* **91**, 015015 (2015).



PORTRAIT OF A MEMORY

Researchers are painting intricate pictures of individual memories and learning how the brain works in the process.

BY HELEN SHEN

For someone who's not a *Sherlock* superfan, cognitive neuroscientist Janice Chen knows the BBC's hit detective drama better than most. With the help of a brain scanner, she spies on what happens inside viewers' heads when they watch the first episode of the series and then describe the plot.

Chen, a researcher at Johns Hopkins University in Baltimore, Maryland, has heard all sorts of variations on an early scene, when a woman flirts with the famously aloof detective in a morgue. Some people find Sherlock Holmes rude while others think he is oblivious to the woman's nervous advances. But Chen and her colleagues found something odd when they scanned viewers' brains: as different people retold their own versions of the same scene, their brains produced remarkably similar patterns of activity¹.

Chen is among a growing number of researchers using brain imaging to identify the activity patterns involved in creating and recalling a specific memory. Powerful technological innovations in human and animal

neuroscience in the past decade are enabling researchers to uncover fundamental rules about how individual memories form, organize and interact with each other. Using techniques for labelling active neurons, for example, teams have located circuits associated with the memory of a painful stimulus in rodents and successfully reactivated those pathways to trigger the memory. And in humans, studies have identified the signatures of particular recollections, which reveal some of the ways that the brain organizes and links memories to aid recollection. Such findings could one day help to reveal why memories fail in old age or disease, or how false memories creep into eyewitness testimony. These insights might also lead to strategies for improved learning and memory.

The work represents a dramatic departure from previous memory research, which identified more general locations and mechanisms. "The results from the rodents and humans are now really coming together," says neuroscientist Sheena Josselyn at the Hospital for Sick Children in Toronto, Canada. "I can't imagine wanting to look at anything else."

ILLUSTRATION BY ANDY POTTS; PHOTOS FROM GETTY

The physical trace of a single memory — also called an engram — has long evaded capture. US psychologist Karl Lashley was one of the first to pursue it and devoted much of his career to the quest. Beginning around 1916, he trained rats to run through a simple maze, and then destroyed a chunk of cortex, the brain's outer surface. Then he put them in the maze again. Often the damaged brain tissue made little difference. Year after year, the physical location of the rats' memories remained elusive. Summing up his ambitious mission in 1950, Lashley wrote²: "I sometimes feel, in reviewing the evidence on the localization of the memory trace, that the necessary conclusion is that learning is just not possible."

Memory, it turns out, is a highly distributed process, not relegated to any one region of the brain. And different types of memory involve different sets of areas. Many structures that are important for memory encoding and retrieval, such as the hippocampus, lie outside the cortex — and Lashley largely missed them. Most neuroscientists now believe that a given experience causes a subset of cells across these regions to fire, change their gene expression, form new connections, and alter the strength of existing ones — changes that collectively store a memory. Recollection, according to current theories, occurs when these neurons fire again and replay the activity patterns associated with past experience.

Scientists have worked out some basic principles of this broad framework. But testing higher-level theories about how groups of neurons store and retrieve specific bits of information is still challenging. Only in the past decade have new techniques for labelling, activating and silencing specific neurons in animals allowed researchers to pinpoint which neurons make up a single memory (see 'Manipulating memory').

IN SEARCH OF THE ENGRAM

Josselyn helped lead this wave of research with some of the earliest studies to capture engram neurons in mice³. In 2009, she and her team boosted the level of a key memory protein called CREB in some cells in the amygdala (an area involved in processing fear), and showed that those neurons were especially likely to fire when mice learnt, and later recalled, a fearful association between an auditory tone and foot shocks. The researchers reasoned that if these CREB-boosted cells were an essential part of the fear engram, then eliminating them would erase the memory associated with the tone and remove the animals' fear of it. So the team used a toxin to kill the neurons with increased CREB levels, and the animals permanently forgot their fear.

A few months later, Alcino Silva's group at the University of California, Los Angeles, achieved similar results, suppressing fear memories in mice by biochemically inhibiting CREB-overproducing neurons⁴. In the process, they also discovered that at any given moment, cells with more CREB are more electrically excitable than their neighbours, which could explain their readiness to record incoming experiences. "In parallel, our labs discovered something completely new — that there are specific rules by which cells become part of the engram," says Silva.

But these types of memory-suppression study sketch out only half of the engram. To prove beyond a doubt that scientists were in fact looking at engrams, they had to produce memories on demand, too. In 2012, Susumu Tonegawa's group at the Massachusetts Institute of Technology in Cambridge reported creating a system that could do just that.

By genetically manipulating brain cells in mice, the researchers could tag firing neurons with a light-sensitive protein. They targeted neurons in the hippocampus, an essential region for memory processing. With the tagging system switched on, the scientists gave the animals a series of foot shocks. Neurons that responded to the shocks churned out the light-responsive protein, allowing researchers to single out cells that constitute the memory. They could then trigger these neurons to fire using laser light, reviving the unpleasant memory for the mice⁵. In a follow-up study, Tonegawa's team placed mice in a new

cage and delivered foot shocks, while at the same time re-activating neurons that formed the engram of a 'safe' cage. When the mice were returned to the safe cage, they froze in fear, showing that the fearful memory was incorrectly associated with a safe place⁶. Work from other groups has shown that a similar technique can be used to tag and then block a given memory^{7,8}.

This collection of work from multiple groups has built a strong case that the physiological trace of a memory — or at least key components of this trace — can be pinned down to specific neurons, says Silva. Still, neurons in one part of the hippocampus or the amygdala are only a tiny part of a fearful foot-shock engram, which involves sights, smells, sounds and countless other sensations. "It's probably in 10–30 different brain regions — that's just a wild guess," says Silva.

A BROADER BRUSH

Advances in brain-imaging technology in humans are giving researchers the ability to zoom out and look at the brain-wide activity that makes up an engram. The most widely used technique, functional magnetic resonance imaging (fMRI), cannot resolve single neurons, but instead shows blobs of activity across different brain areas. Conventionally, fMRI has been used to pick out regions that respond most strongly to various tasks. But in recent years, powerful analyses have

revealed the distinctive patterns, or signatures, of brain-wide activity that appear when people recall particular experiences. "It's one of the most important revolutions in cognitive neuroscience," says Michael Kahana, a neuroscientist at the University of Pennsylvania in Philadelphia.

The development of a technique called multi-voxel pattern analysis (MVPA) has catalysed this revolution. Sometimes called brain decoding, the statistical method typically feeds fMRI data into a computer algorithm that auto-

matically learns the neural patterns associated with specific thoughts or experiences. As a graduate student in 2005, Sean Polyn — now a neuroscientist at Vanderbilt University in Nashville, Tennessee — helped lead a seminal study applying MVPA to human memory for the first time⁹. In his experiment, volunteers studied pictures of famous people, locations and common objects. Using fMRI data collected during this period, the researchers trained a computer program to identify activity patterns associated with studying each of these categories.

Later, as subjects lay in the scanner and listed all the items that they could remember, the category-specific neural signatures re-appeared a few seconds before each response. Before naming a celebrity, for instance, the 'celebrity-like' activity pattern emerged, including activation of an area of the cortex that processes faces. It was some of the first direct evidence that when people retrieve a specific memory, their brain revisits the state it was in when it encoded that information. "It was a very important paper," says Chen. "I definitely consider my own work a direct descendant."

Chen and others have since refined their techniques to decode memories with increasing precision. In the case of Chen's *Sherlock* studies, her group found that patterns of brain activity across 50 scenes of the opening episode could be clearly distinguished from one another. These patterns were remarkably specific, at times telling apart scenes that did or didn't include Sherlock, and those that occurred indoors or outdoors.

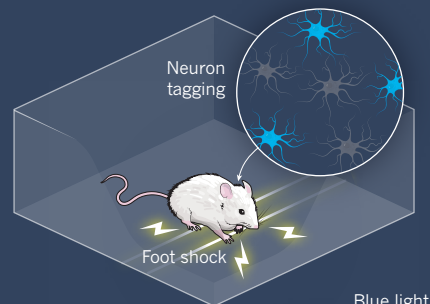
Near the hippocampus and in several high-level processing centres such as the posterior medial cortex, the researchers saw the same scene-viewing patterns unfold as each person later recounted the episode — even if people described specific scenes differently¹. They even observed similar brain activity in people who had never seen the show but had heard others' accounts of it¹⁰.

"It was a surprise that we see that same fingerprint when different people are remembering the same scene, describing it in their own words, remembering it in whatever way they want to remember," says

**"I CAN'T IMAGINE
WANTING TO LOOK AT
ANYTHING ELSE."**

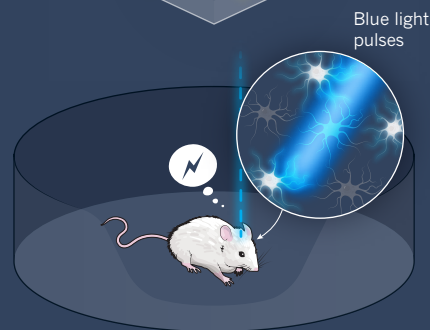
MANIPULATING MEMORY

To identify neurons that form part of a memory engram, researchers have developed systems for tagging, reactivating and silencing them.



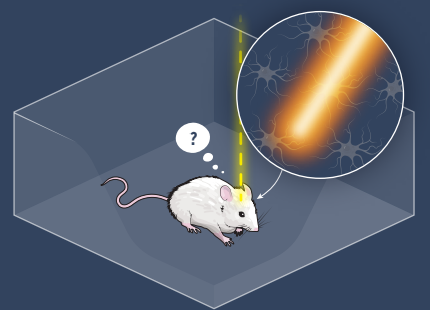
NEURON TAGGING

Cells in the hippocampus are altered so that when they fire, they produce a light-sensitive protein. The mouse forms a memory of a shock to the foot, and the neurons that are activated are tagged.



MEMORY RECALLED

Researchers can induce the tagged neurons to fire using a blue laser. Even in a different cage, the mouse recalls the foot shock.



MEMORY SUPPRESSED

To block a memory, some studies use a protein that silences cells when exposed to light of a certain colour. Even in the cage where it formed the foot-shock memory, the mouse cannot retrieve it.

Chen. The results suggest that brains — even in higher-order regions that process memory, concepts and complex cognition — may be organized more similarly across people than expected.

MELDING MEMORIES

As new techniques provide a glimpse of the engram, researchers can begin studying not only how individual memories form, but how memories interact with each other and change over time.

At New York University, neuroscientist Lila Davachi is using MVPA to study how the brain sorts memories that share overlapping content. In a 2017 study with Alexa Tompar, then a graduate student in her lab, Davachi showed volunteers pictures of 128 objects, each paired with one of four scenes — a beach scene appeared with a mug, for example, and then a keyboard; a cityscape was paired with an umbrella, and so on. Each object appeared with only one scene, but many different objects appeared with the same scene¹¹. At first, when the volunteers matched the objects to their corresponding scenes, each object elicited a different brain-activation pattern. But one week later, neural patterns during this recall task had become more similar for objects paired with the same scene. The brain had reorganized memories according to their shared scene information. “That clustering could represent the beginnings of learning the ‘gist’ of information,” says Davachi.

Clustering related memories could also help people use prior knowledge to learn new things, according to research by neuroscientist Alison Preston at the University of Texas at Austin. In a 2012 study,

Preston’s group found that when some people view one pair of images (such as a basketball and a horse), and later see another pair (such as a horse and a lake) that shares a common item, their brains reactivate the pattern associated with the first pair¹². This reactivation appears to bind together those related image pairs; people that showed this effect during learning were better at recognizing a connection later — implied, but never seen — between the two pictures that did not appear together (in this case, the basketball and the lake). “The brain is making connections, representing information and knowledge that is beyond our direct observation,” explains Preston. This process could help with a number of everyday activities, such as navigating an unfamiliar environment by inferring spatial relationships between a few known landmarks. Being able to connect related bits of information to form new ideas could also be important for creativity, or imagining future scenarios.

In a follow-up study, Preston has started to probe the mechanism behind memory linking, and has found that related memories can merge into a single representation, especially if the memories are acquired in close succession¹³. In a remarkable convergence, Silva’s work has also found that mice tend to link two memories formed closely in time. In 2016, his group observed that when mice learnt to fear foot shocks in one cage, they also began expressing fear towards a harmless cage they had visited a few hours earlier¹⁴. The researchers showed that neurons encoding one memory remained more excitable for at least five hours after learning, creating a window in which a partially overlapping engram might form. Indeed, when they labelled active neurons, Silva’s team found that many cells participated in both cage memories.

These findings suggest some of the neurobiological mechanisms that link individual memories into more general ideas about the world. “Our memory is not just pockets and islands of information,” says Josselyn. “We actually build concepts, and we link things together that have common threads between them.” The cost of this flexibility, however, could be the formation of false or faulty memories: Silva’s mice became scared of a harmless cage because their memory of it was formed so close in time to a fearful memory of a different cage. Extrapolating single experiences into abstract concepts and new ideas risks losing some detail of the individual memories. And as people retrieve individual memories, these might become linked or muddled. “Memory is not a stable phenomenon,” says Preston.

Researchers now want to explore how specific recollections evolve with time, and how they might be remodelled, distorted or even recreated when they are retrieved. And with the ability to identify and manipulate individual engram neurons in animals, scientists hope to bolster their theories about how cells store and serve up information — theories that have been difficult to test. “These theories are old and really intuitive, but we really didn’t know the mechanisms behind them,” says Preston. In particular, by pinpointing individual neurons that are essential for given memories, scientists can study in greater detail the cellular processes by which key neurons acquire, retrieve and lose information. “We’re sort of in a golden age right now,” says Josselyn. “We have all this technology to ask some very old questions.”■

Helen Shen is a science journalist based in Sunnyvale, California.

1. Chen, J. *et al. Nature Neurosci.* **20**, 115–125 (2016).
2. Lashley, K. S. *Soc. Exp. Biol. Symp.* **4**, 454–482 (1950).
3. Han, J. H. *et al. Science* **323**, 1492–1496 (2009).
4. Zhou, Y. *et al. Nature Neurosci.* **12**, 1438–1443 (2009).
5. Liu, X. *et al. Nature* **484**, 381–385 (2012).
6. Ramirez S. *et al. Science* **341**, 387–391 (2013).
7. Denny, C. A. *et al. Neuron* **83**, 189–201 (2014).
8. Tanaka, K. Z. *et al. Neuron* **84**, 347–354 (2014).
9. Polyn, S. M., Natu, V. S., Cohen, J. D. & Norman, K. A. *Science* **310**, 1963–1966 (2005).
10. Zádbood, A., Chen, J., Leong, Y. C., Norman, K. A. & Hasson, U. *Cereb. Cortex* **27**, 4988–5000 (2017).
11. Tompar, A. & Davachi, L. *Neuron* **96**, 228–241 (2017).
12. Zeithamova, D., Dominick, A. L. & Preston, A. R. *Neuron* **75**, 168–179 (2012).
13. Zeithamova, D. & Preston, A. R. *J. Cogn. Neurosci.* **29**, 1311–1323 (2017).
14. Cai, D. J. *et al. Nature* **534**, 115–118 (2016).

COMMENT

SUSTAINABILITY Two very different visions of how to secure the future **p.152**



HISTORY Did science help women to get the vote? **p.154**

PUBLISHING Hold an annual test to highlight hoax journals **p.155**

CLIMATE CHANGE Study fragile ecosystems' response to warming **p.155**

NASA/JPL/UNIV. ARIZONA



Impact craters and atmospheric history on Mars provide information on how terrestrial planets form and evolve.

Exoplanet science 2.0

The study of life on and off Earth needs unified funding and a coherent plan, say **Caleb Scharf, Debra Fischer and Victoria Meadows.**

It is more than two decades since we learnt that the Universe is awash with other worlds. Since 1992, more than 3,500 exoplanets have been discovered orbiting stars other than our Sun.

The range of systems is dazzling. There is at least one planet around any star that, like the Sun, is powered by fusing hydrogen into helium. Sixty per cent of such stars harbour 'super-Earths' — rocky worlds that are more massive than ours but smaller than

Neptune. One in six of these stars has an Earth-sized planet in an orbit that is tighter than Mercury's around the Sun¹.

This plethora of rocky planets raises a big question: is life common in the Universe? Even in our Solar System, there are plenty of places where organisms could potentially survive, such as in the oceans of liquid water beneath the frozen surfaces of Jupiter's satellite Europa and Saturn's moon Enceladus. Four billion years ago,

life may have thrived on a warmer Mars.

Within a decade or two, we might find traces of extraterrestrial life in our Solar System. The Mars 2020 and ExoMars 2020 rovers are set to probe the Martian surface in that year. NASA's Europa Clipper and the European Space Agency's Jupiter Icy Moons Explorer (JUICE) ventures will get close to Jupiter's satellites by about 2030. The James Webb Space Telescope will look farther afield, scrutinizing the atmospheres of ►

► distant exoplanets in deep space².

Insights from many disciplines are needed to discover which ingredients, mechanisms and environmental pathways create and sustain life. Molecular biologists need to explain how proto-life might operate. Evolutionary biologists and ecologists need to probe life's interplay with alien environments. Geophysicists, geochemists and planetary scientists need to describe how planets evolve over billions of years. And astronomers have to detect more remote biospheres, while astrobiologists help to tie the pieces together.

Exoplanetary exploration should be central to this quest. Although exoplanets pique public attention, some astronomers see this field as niche and immature — they prefer to leave the review and funding of interdisciplinary projects in exoplanetary science to other fields. But if astronomers aren't included in such efforts, scientific quality suffers. Exoplanet science requires large and expensive teams, telescopes, satellites and computing facilities. But allied fields such as planetary and Earth science are established, vibrant and have their own wish lists of discipline-specific projects that are more ready for action than those in exoplanet research.

Competition over resources and intellectual turf is fierce among all these fields. For example, astronomers may favour building space-based observatories to gather more statistical data on exoplanets³. Meanwhile, planetary scientists might argue for detailed studies of a few planets. Both approaches are ultimately compatible, but that tension erodes the clarity of goals and can make funders nervous.

Crucial opportunities for scientists to learn from one another are falling between the cracks. For example, most Solar-System research is barely influenced by exoplanetary studies, and vice versa. Yet exoplanet data must be calibrated with knowledge about the Solar System, from the nature of runaway greenhouse-gas effects on Venus-like planets to how the orbits of young planetary systems are reconfigured.

INTERACTION, NOT ISOLATION

There has to be a radical shift. Now that answers about life's universality are finally within reach, funding agencies and scientists must step up. In our view, the field needs a systems-science approach⁴ focused on interactions — between galactic environments, planet formation, orbital dynamics, heliophysics, atmospheres, hydrospheres, cryospheres, geospheres, biospheres and magnetospheres — rather than on components in isolation. This would extend Earth-systems science to encompass other types of planet and ecosystem.

Here we highlight three key questions that



Studying organisms from Yellowstone National Park's hot springs can uncover conditions needed for life.

illustrate how exoplanet systems science can draw disciplines together.

What dictates planets' variety and properties? For example, why are the atmospheres and climates of Venus, Earth, Mars and Titan so different? To find out, we must bridge the gaps between Solar-System, exoplanet and astrophysical science. Observational data must be tied to models that simulate the evolution of the atmospheres, interiors and surfaces of planets over billions of years⁵. Tools from data science must be adapted to tackle increasingly large and complex data sets.

The Solar System should serve as one calibration point while its statistical significance is assessed. For example, structures in Jupiter's atmosphere and magnetic field revealed by NASA's Juno spacecraft are changing views of the planet's core and of how gas giants form. Studies of vortices and reflective particles in Neptune's atmosphere have shown how chemistry affects the spectra of ice giants. And the New Horizons mission to the dwarf planet Pluto and the Dawn mission to the minor planets Vesta and Ceres helped to trace how condensed volatile compounds are distributed in the Solar System.

Exoplanetary data challenge established ideas and put our understanding of the Solar System into a wider context. For example, we now know that planets can form around binary stars, extremely close to stars and in dense packs. Gas giants have a wider range of chemical compositions than was previously thought. Planetary orbits can be highly elongated or inclined. Astronomy facilities such as the Atacama Large Millimeter/submillimeter Array (ALMA) in Chile are revealing details of the agglomeration of dust and solids,

and chemical zones in nascent planetary systems unlike ours.


Wider insights from astronomy are also needed. A major question is how stars influence the planets around them. Stars spin and oscillate according to their age, internal structure and activity. Young and low-mass stars can emit intense X-rays and γ -rays or eject charged particles. These may erode the atmospheres of planets and modify their composition, affecting their surface temperature and ability to hold water⁶. A planet's magnetosphere can mitigate this, but needs to be better understood.

The elements in stars influence planet formation, but it is unclear how. Elements can accumulate in different areas of the disks that ring young stars. The build-up of material might be affected by the rates at which stars and disks spin. The bulk properties of stars and their births across the Milky Way need to be investigated in more depth to establish how planets have formed from the Big Bang to today.

How can we identify worlds that are capable of harbouring life?

The study of exoplanets opens up a wider range of planetary characteristics than we can observe in the Solar System alone, such as mass, composition and orbital configuration. Knowledge of Earth's deep environmental history, climate and chemical state is essential for calibrating models that explore the likelihood of life forming on other worlds, perhaps under different conditions. But a broader approach to planets would also help to interpret Earth: from the puzzles of ancient atmospheric oxygenation and chemical and climatic change, to the influence of human activity.

Geoscientists and astronomers need to develop better criteria for categorizing



planets, including those capable of hosting life. Concepts such as the ‘habitable zone’ around stars can guide our initial search, by simplistically identifying rocky planets that might have liquid water on the surface. But the real challenge lies in modelling and measuring actual details of surface conditions and imagining evolutionary strategies in these places⁷. The presence of temperate surfaces depends on many things, including the composition and photochemistry of the atmosphere, the tilt and rate at which a planet spins and the topography of a planet’s surface⁵. A systems approach would be much more efficient at formally identifying the most important factors than current methods are.

Existing efforts that bring climate scientists together with astronomers to build generalized climate models for rocky exoplanets could be the kernel for growing this systems approach. These models, in turn, test the sensitivity of Earth’s properties to atmospheric conditions and extreme forcings of climate.

Basic geological research is needed to understand the cores of planets, the weathering and transport of material on their surfaces, their magnetic fields and the probability that water is present. Exoplanetary science is stimulating advances in deep-Earth sensing, experimentation and modelling⁸. For example, the 2017 American Geophysical Union (AGU) autumn meeting hosted sessions on how heat and volcanism influence the geochemistry, mineralogy and petrology of Mercury, Venus, Earth, the Moon, Mars and asteroids.

How can we decode life’s relationship with its environment? Life’s possible behaviour on planets around other stars with different orbits, ages and histories is central to understanding Earth systems and the origins and early evolution of life on our planet. Microbiologists and astrobiologists need to inform speculations about life elsewhere by providing limits to its molecular capabilities. It is helpful to study terrestrial organisms that live in extreme conditions, such as around deep-sea hydrothermal vents or hot springs, but astronomers and planet modellers must know the options for life’s possible effects on planetary chemistry and its interplay with abiotic processes if they are to find it. Work on metabolic pathways and on abiotic photochemistry and geochemistry is changing perspectives on chemical biomarkers and global chemical equilibria⁹.

We need to know what fraction of a planet is capable of sustaining organisms, as well as which chemical and climatic properties that can be observed astronomically may reveal a biosphere. Ecological models in Earth-climate simulations need to be examined in the context of exoplanets, where radiation, rotation, planet orientation

and land–ocean fractions are very different. Fundamental questions about cell function and adaptation can be tackled theoretically and experimentally using virtual and laboratory environments. Ecologists, planetary scientists and geoscientists must also examine the nature of geospheres for planets of widely different ages, as well as primitive atmospheres where molecular species such as hydrogen may be abundant.

Uncertainties about the chemical and thermal conditions of young planets must be reduced. Where do the first biomolecules come from, and what chemistry is involved in life’s origins? Data from exoplanetary systems, as well as from laboratory astrochemistry and models of planet assembly, can provide scenarios for chemists and biologists to evaluate and study these processes experimentally.

NEW FRONTIERS

Exoplanetary systems science will be kick-started through the reorientation of research and the restructuring of funding programmes. Funding agencies should replace current grant silos with broader themes. For example, elements of the US National Science Foundation’s (NSF’s) Astronomy & Astrophysics, Geophysics and Ecosystem Studies programmes could be replaced by one exoplanetary systems science programme.

The NSF’s solar and planetary research programme, NASA’s Cosmic Origins programme and the European Research Council’s Synergy Grant scheme still largely assign funding in traditional ways. Fields such as Solar-System science and exoplanetary science should not have to compete. It is essential that agencies and institutions support systems-inspired consortia.

The next-generation of space-based observatories that are being discussed for selection in 2020 and launch in the 2030s should be viewed as systems-science missions. These include NASA’s Large UV/Optical/IR Surveyor (LUVOIR) or Habitable Exoplanet Imaging Mission (HabEx). Their priorities should be evaluated in an interdisciplinary light and plans should be made accordingly for how their time will be allocated¹⁰.

Some institutions have already moved in this direction. Since 1998, the NASA Astrobiology Institute, directed from NASA’s Ames Research Center in Mountain View, California, has funded astrophysics, exoplanets, biology, chemistry and planetary exploration through a single programme. Some universities, such as the University of Arizona in Tucson, the University of Washington in Seattle and McMaster University in Hamilton, Canada, have established

centres and graduate programmes that bridge astronomy, planetary science, Earth science and biological sciences.

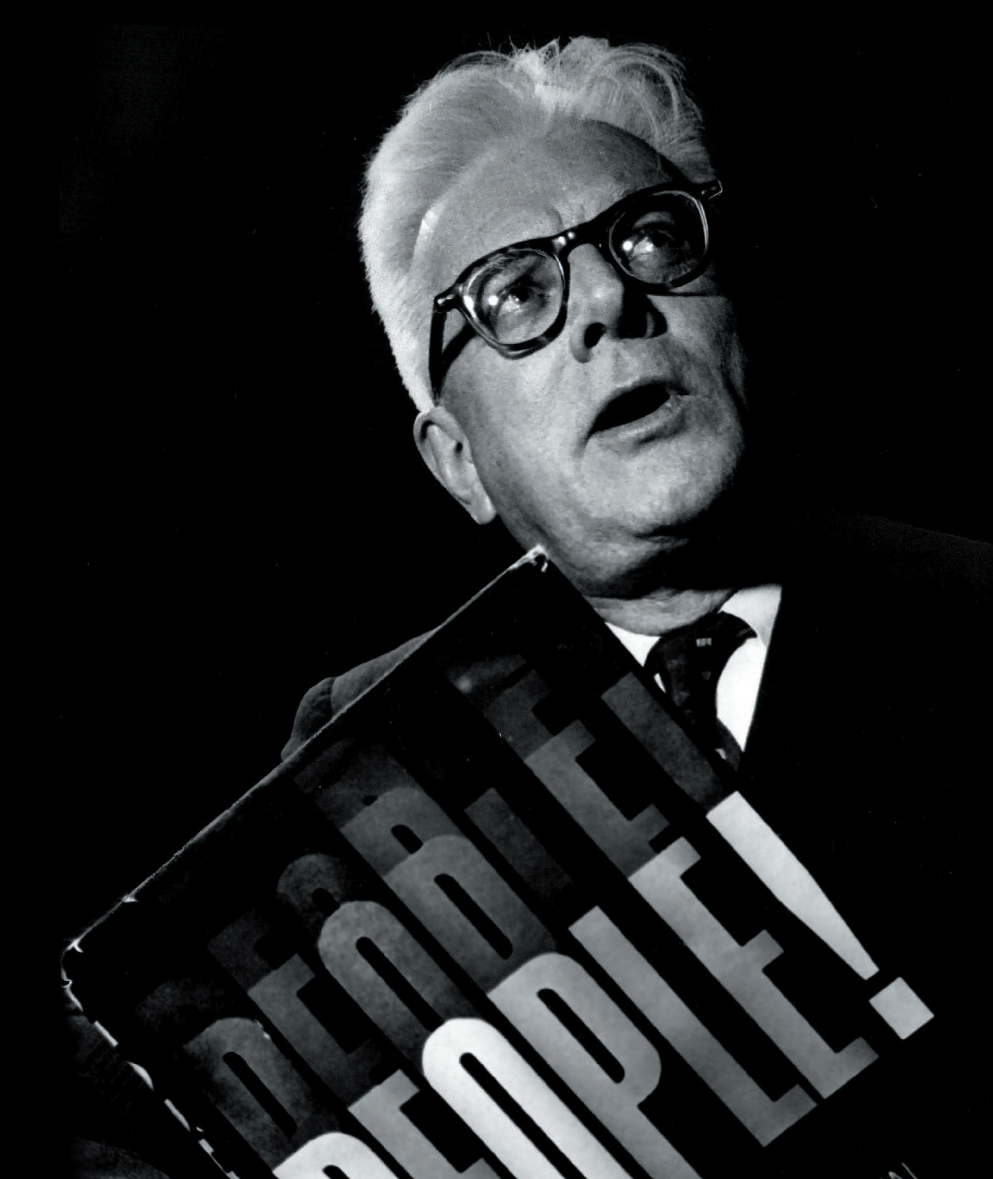
Networks are being created, such as the European Astrobiology Campus and the European Astrobiology Network Association, to foster interdisciplinary training and communication. Efforts are under way to accelerate astrobiology research in China, initiated by a team formed at the International Space Science Institute in Bern, Switzerland. Since 2015, NASA’s Nexus for Exoplanetary System Science (NExSS) coalition has forged a community that supports the exchange of ideas and active collaboration. It comprises more than a dozen teams with diverse approaches to modelling and observing exoplanets.

Building more coherence into efforts such as these would be the next step towards exoplanetary systems science. It must be the subject of a bigger conversation before the next US decadal surveys, in 2020 for astronomy and in 2022 for planetary science. We encourage professional societies to address the idea. These include the American Astronomical Society, the AGU and the American Association for the Advancement of Science (AAAS) and global organizations such as the International Astronomical Union (IAU).

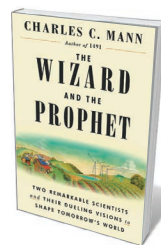
A good start would be for the AAAS or the IAU to convene researchers from areas that are already embracing systems approaches to share their insights with exoplanetary researchers. We have a lot to learn from genomics, systems biology, complex systems, public health, data science and machine learning. ■

Caleb Scharf is director of astrobiology at Columbia University, New York City, New York, USA. **Debra Fischer** is a professor of astronomy at Yale University, New Haven, Connecticut, USA. **Victoria Meadows** is a professor of astronomy and principal investigator at the Virtual Planet Laboratory, University of Washington, Seattle, Washington, USA.
e-mail: caleb@astro.columbia.edu

1. Dressing, C. D. & Charbonneau, D. *Astrophys. J.* **807**, 45 (2015).
2. Deming, D. et al. *Publ. Astron. Soc. Pac.* **121**, 952–967 (2009).
3. Bean, J. L., Abbot, D. S. & Kempton, E. M.-R. *Astrophys. J. Lett.* **841**, L24 (2017).
4. Mobus, G. E. & Kalton, M. C. *Principles of Systems Science* (Springer, 2015).
5. Mackwell, S. J. et al. (eds) *Comparative Climatology of Terrestrial Planets* (Univ. Arizona Press, 2013).
6. Zahnle, K. J. & Catling, D. C. *Astrophys. J.* **843**, 122 (2017).
7. Kasting, J. F. & Catling, D. *Annu. Rev. Astron. Astrophys.* **41**, 429–463 (2003).
8. Duffy, T. S., Madhusudhan, N. & Lee, K. K. M. in *Mineralogy of Super-Earth Planets. Treatise on Geophysics* 149–178 (Elsevier, 2015).
9. Meadows, V. S. *Astrobiology* **17**, 1022–1052 (2017).
10. Fujii, Y. et al. Preprint at <https://arxiv.org/abs/1705.07098> (2017).



Ecologist William Vogt warned of the dangers of dwindling resources.



The Wizard and the Prophet: Two Remarkable Scientists and Their Dueling Visions to Shape Tomorrow's World
CHARLES C. MANN
Knopf: 2018.

DENVER POST VIA GETTY

SUSTAINABILITY

Duel for the future

Adam Rome assesses a study of two scientists who have polarized attitudes to sustainability since the 1960s.

Our species has had an amazingly successful run. Billions of people now live in environments radically transformed to suit human needs and wants. But humanity's future is far from guaranteed. How will we meet the looming challenges of the twenty-first century? We can work even harder to master the planet with technological ingenuity. Or we might need to accept that our desires can't be unlimited, and see ourselves as citizens of a larger-than-human community, rather than as world conquerors. We can't do both, science writer Charles Mann argues in *The Wizard and the Prophet*, an effort to assess which path holds the more promise.

To dramatize the two options, Mann

contrasts the work of agronomist Norman Borlaug (the Wizard of his title) with that of ecologist William Vogt (the Prophet). In 1970, Borlaug won the Nobel Peace Prize for developing high-yield varieties of wheat that launched the Green Revolution. Along with agricultural chemicals and irrigation systems, Borlaug's seeds led to a sharp rise in productivity in Mexico, India and other developing countries, particularly in the 1960s. Vogt's 1948 best-seller *Road to Survival* warned that rising population and declining resources spelt global catastrophe. Whereas Borlaug hoped to free humanity from the constraints of nature, Vogt called for a new environmental consciousness.

Although few today would self-identify as followers of Borlaug or Vogt, the heart of Mann's book asks how people he considers their intellectual heirs propose to deal with climate change and to provide food, water and energy for a projected global population of 10 billion (or more) by 2050. His Wizard camp ranges from biotech boosters to advocates of geoengineering. His Prophets include the authors of *The Limits to Growth* (Universe, 1972), along with the small-is-beautiful advocates of organic agriculture and solar power.

The structure of *The Wizard and the Prophet* reminded me of John McPhee's brilliant *Encounters with the Archdruid* (Farrar, Straus and Giroux, 1971). That book explored the implications of the environmental movement by arranging confrontations between David Brower — long-time leader of the conservationist Sierra Club and founder of Friends of the Earth — and three presumed foes. Brower debates a mining engineer, a resort developer and a dam builder (the latter, on a raft trip on a wild stretch of the Colorado River). McPhee respected all four, and was masterful at challenging stereotypes. Readers were free to decide who had won the debates.

Unfortunately, Mann's study doesn't measure up to McPhee's classic. It is flawed in many ways, most notably in its lack of even-handedness. Mann writes that he was a Vogtian when young, later became a Borlaugian and is now torn — but I don't see that ambivalence in the text. Mann indicts Vogt as a failure who wasted precious time by leading people down a dead end. He considers Borlaug a saviour, even though the Green Revolution had unfortunate social and environmental consequences, such as a growing concentration of land ownership and pollution of waterways through overuse of pesticides. Mann also stacks the deck by ignoring problems with the Borlaugian approach and neglecting compelling elements of the Prophetic tradition.

At root, the differences between Borlaug and Vogt were ideological, not scientific. Borlaug accepted the mainstream values of his time and place — the American dream of material progress. Vogt didn't; like all prophets, he was a critic. He called for people to reappraise their place in the world: to think ecologically about everything from what we consume to how we understand history. He questioned whether "that sacred cow Free Enterprise" could be environmentally sustainable. And he advocated population

control, which went against many people's religious views and humanist ideals. Aside from decrying the latter notion, Mann engages with none of these ideas.

Instead, Mann turns the ideological divide into a dispute about technological visions, the hard and soft paths (a dichotomy he appropriates from physicist Amory Lovins). Wizards favour 'hard', sophisticated, capital-intensive, top-down methods of ensuring adequate food, water and energy, Mann argues. Prophets believe in simpler, decentralized, 'soft' solutions. But that definition is Borlaugian. It assumes that the goal is to meet ever-greater demand for natural resources — a premise that most Vogtians reject, because they argue that we need to moderate our desires, not just find less destructive ways to slake them. Even if Mann considers that argument naive, fairness demands giving it a hearing.

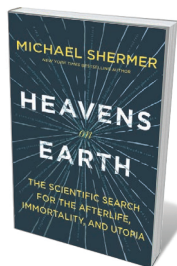
Mann also caricatures proponents of the soft path — particularly Lovins. Lovins is as can-do as any techhead; he's not a counter-cultural guru. Yet he does warn in *Soft Energy Paths* (Friends of the Earth International, 1977) that hard technologies lead to undemocratic concentrations of power, as major oil companies have proved. He is also a leader in making the market greener, as a consultant to corporations and as co-author of *Natural Capitalism* (Little, Brown, 1999). Although Mann dismisses him as a retro activist, Lovins would be a worthy antagonist for any Borlaugian.

And it's to the Borlaugians that Mann is most generous. He considers the evidence for the safety of genetically engineered crops as compelling as the scientific consensus on climate change. He holds out hope for nuclear power. And he barely acknowledges that history provides countless reasons for anxiety about unintended consequences of technology. From plastics to chemical pesticides, many twentieth-century miracles have done harm as well as good. Even some technology boosters admit that surprises are inevitable, although they remain undaunted. As the automotive pioneer Charles Kettering liked to say: "The price of progress is trouble, and I don't think the price is too high."

Mann asserts that those who lean towards Vogt's world view can't prove that we'll hit planetary limits. But the heirs of Borlaug can't prove that they'll avoid making a mistake that undermines the ecological or planetary foundations of civilization. Where does that leave us? The Wizards have had most of the momentum since the Enlightenment. The Prophets keep the Wizards from overreaching, and challenge us to probe what we really value. We need to listen carefully to both. ■

Adam Rome is a professor of history at the University at Buffalo, New York. His latest book is *Green Capitalism?*.
e-mail: adamrome@buffalo.edu

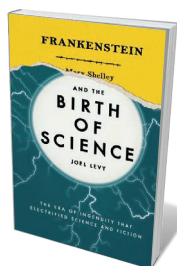
Books in brief



Heavens on Earth

Michael Shermer HENRY HOLT (2018)

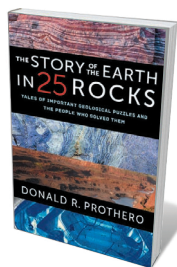
An astonishing 75% of US citizens — including some avowed atheists — believe in an afterlife. So potent is the idea of immortality, reminds *Skeptic* magazine publisher Michael Shermer in this intriguing study, that it pervades human culture. After exploring the notion's place in religious belief, Shermer examines its scientific manifestations, from transhumanism and longevity research to cryonics. He looks, too, at utopianism as the desire to create an earthly paradise. He concludes that balanced rationality — along with an honest, positive acceptance of mortality — constitutes the real "soul" of life.



Frankenstein and the Birth of Science

Joel Levy ANDRE DEUTSCH (2018)

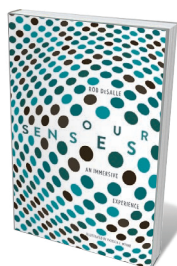
The bicentennial of Mary Shelley's masterwork *Frankenstein* is upon us. And one of the first homages of the year is this episodic, entertaining analysis by science writer Joel Levy. He presents the novel as a portrayal of high-Romantic "gonzo science", as well as science fiction. Levy contextualizes Shelley's narrative with contemporary research into areas such as galvanic revivification, psychoactive substances and polar discovery (as Victor Frankenstein and his monster travel to the North Pole). A celebration of an enduring classic's "extraordinarily rich confluence of sources".



The Story of the Earth in 25 Rocks

Donald R. Prothero COLUMBIA UNIVERSITY PRESS (2018)

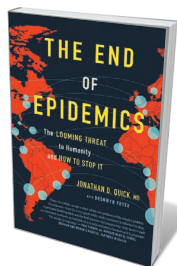
Geologist Donald Prothero has crafted a rock-solid premise for this delightful book: a tour of 25 geological discoveries that changed our understanding of Earth and the cosmos. He begins explosively, with Pliny the Younger's eyewitness account of the eruption of Vesuvius in southern Italy in AD 79 — the first scientifically accurate description of such an event. He then reveals how deep time, the Moon's origins and other 'stories in stones' were cracked by luminaries from Enlightenment geologist James Hutton to Marie Tharp, who mapped the Atlantic Ocean's floor in the 1950s.



Our Senses

Rob DeSalle YALE UNIVERSITY PRESS (2018)

Sight, hearing, touch, smell, taste: the senses are our portal to the world. But this erudite, zesty study by Rob DeSalle, curator at the American Museum of Natural History in New York City, ranges far beyond these "big five" into arenas such as balance, pain, heat and cold. DeSalle examines sense in an array of fauna, including comb jellies, lampreys and bats. He digs deepest, however, into how perception is formed in the human brain, how phenomena such as synaesthesia arise, how people with brain damage experience the world, and how our sensory armoury feeds creativity.



The End of Epidemics

Jonathan D. Quick and Bronwyn Fryer ST MARTIN'S PRESS (2018)

Physician Jonathan Quick's long experience at the front lines of global public health gives his call to action on pandemics a searing urgency. With writer Bronwyn Fryer, Quick examines how fear and complacency impede responses to emergencies such as the 2014 Ebola epidemic in West Africa. He then sets out a seven-part solution centred on actions such as establishing resilient health systems and mobilizing on-the-ground activism. Pragmatic, insightful and research-rich, this is a key volume for the policymaker's shelf. **Barbara Kiser**



Botanist Helen Gwynne-Vaughan was controller of the British Women's Army Auxiliary Corps.

HISTORY

Science and suffrage

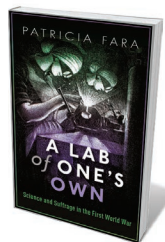
Elizabeth Bruton lauds a book tracing how women in wartime research blazed a path to the vote and beyond.

A century ago, women over 30 were granted the vote in Britain. (US women gained the vote two years later, although African Americans and Native Americans were still effectively disenfranchised for some years.) The UK watershed coincided with the end of the First World War. Historian of science Patricia Fara commemorates the moment with *A Lab of One's Own*, using archival research to draw together narratives of science, war and suffrage (as she trailed in an essay: *Nature* **511**, 25–27; 2014).

The standard take on this period is that British women gained opportunities through labour shortages, the result of 6 million men going to war. Thus, women were able to enter fields such as science, technology, engineering, mathematics and medicine (STEMM).

Fara's story differs. She shows how women's entry into these areas was shaped by the prewar efforts and example of exceptional women including archaeologist Agnes Conway; biochemist Ida Smedley; and political campaigner Ray Strachey, related to Virginia Woolf. (The title of Fara's book, suggested by historian Marsha Richmond, was inspired by Woolf's classic 1929 'A Room of One's Own'.)

Along with agitating for the vote, these women called for more than the traditional roles of domesticity, clerical work, nursing and teaching. They lobbied for professional opportunities, financial independence and



A Lab of One's Own: Science and Suffrage in the First World War
PATRICIA FARA
Oxford University Press: 2018.

Carl Linnaeus, Fara claims, used their theories to argue for the impossibility of sexual equality. In 1904, chemist Henry Armstrong argued that, because women were thought to be lower down the evolutionary scale, "education can do little" to modify their nature.

Fara's nuanced narrative centres on a group of scientific and medical women, many of them graduates of Newnham College, Cambridge. Strachey studied mathematics before turning to politics, fighting for women's economic, professional and political power before, during and after the war. Conway studied history and chronicled women's work. Smedley was the first woman

higher degrees. Fara shows how they created opportunities in research, medicine, intelligence and code-breaking. They opened doors in factories, academia, hospitals and the battlefield.

They also fought the belief that women were inherently lesser than men, shaped by biological justifications, including eugenics. Charles Darwin and founder of taxonomy

admitted to the London Chemical Society.

Among the non-Newnhamites, Caroline Haslett rose from the post of clerk at the Cochran Boiler Company (which made parts for ships) to train as an engineer during the war. Later, she became the first female member of the British Electricity Authority. Formidable Scottish geologist Maria Gordon was the first woman to be awarded a doctor of science from the University of London, in 1893. This group is completed by the "scientists in khaki" and leaders of the Women's Army Auxiliary Corps, physician Mona Geddes and botanist Helen Gwynne-Vaughan.

Fara also highlights achievements of lesser-known women. We meet aeronautical researcher Beatrice Mabel Cave-Browne-Cave; spycatcher Mabel Elliott; and the diplomatic-mail readers of the Admiralty's Room 40 who, with their codebreaking counterparts, saw their covert wartime work persist into peacetime. Fara discusses, too, medical luminaries such as Helena Gleichen and Nina Hollings, who worked in new fields including radiography and physiotherapy. Interwoven are fascinating glimpses of women about whom "only snippets of information" survive. Fara's retrieval of them makes this narrative more than the sum of its parts.

But winning the war, and the vote, did not result in equality: it would be another decade before the Equal Franchise Act of 1928 granted voting parity. And the interwar years saw a return to prewar mores. Male veterans reclaimed jobs, and women's opportunities dried up, among expectations that they would return to the kitchen.

If there is a weakness in Fara's approach, it is that the focus on Cambridge graduates veers close to a 'Great Women' echo of the 'Great Men' history that Fara criticizes. She does acknowledge, if sparsely, difficulties experienced by working-class women, for example in gas production and munitions. Nevertheless, she shows how women and their wartime work changed perceptions of female roles and competency, and influenced professional and educated women earning their own living. In 1919, the Women's Engineering Society was founded. A year later, the University of Oxford granted women the right to graduate.

The wartime changes were neither long-standing nor wide-ranging. But they were — Fara argues — catalysts for many positive shifts in the workplace. The discrimination experienced by many of the women in *A Lab of One's Own* is now illegal. Fara concludes with an open-ended question: how can what we learn from this history challenge other historical interpretations, and so inform the future narratives of women in STEMM? ■

Elizabeth Bruton is curator of technology and engineering at London's Science Museum. e-mail: elizabeth.bruton@sciencemuseum.ac.uk

Correspondence

Gene-drive e-mails legally requested

Gene-drive technology does indeed need proper scrutiny, but it also needs transparent and accountable governance (see *Nature* **552**, 6; 2017). In my view, your Editorial seems to be trying to excuse the influence of big-money manipulations on scientific decision-making when it comes to this risky technology. In so doing, I feel that *Nature* has crossed a line in conflating a cornerstone of investigative journalism — requests under freedom-of-information laws — with outright theft.

Specifically, you write that the release of 1,200 e-mails from gene-drive researchers (obtained by Edward Hammond under US open-records laws) “echoes the way in which hackers released documents stolen from climate scientists before a major UN meeting in 2009”. The two incidents are very different. Those ‘Climategate’ e-mails were taken illegally. These e-mails, dubbed the Gene Drive Files, were released by the institutions involved in accordance with legal requirements. There is nothing criminal about this.

Furthermore, climate deniers used the Climategate e-mails to claim that data had been falsified. Scientists robustly and correctly responded through independent inquiries that this was incorrect. By contrast, the Gene Drive Files concern issues of process: they corroborate how the Bill & Melinda Gates Foundation in Seattle, Washington, paid US\$1.6 million to a private public-relations firm, apparently with the intention of influencing the United Nations discussion on gene drives by coordinating what an ‘advocacy coalition’ of public researchers should say in an expert process. *Nature* failed to provide the details but readers can make up their own minds.

Employment as a researcher at a publicly funded institution is an immense privilege. Such researchers are rightly

accountable to the public — not to private public-relations firms or big-money agendas. Accountability is exactly why we have freedom-of-information laws. Undermining those laws undermines a free press.

Jim Thomas *ETC Group, Montreal, Canada.*
jim@etcgroup.org
J.T. declares competing financial interests; see go.nature.com/2ctjftu

Arm against return of breast cancer

Your summary of the latest study by the Early Breast Cancer Trialists’ Collaborative Group, states that “Even after treatment, odds of recurrence are worse for the next 20 years” (*Nature* <http://go.nature.com/2eob74j>; 2017). We find this statement unnecessarily alarming.

The same group showed in previous work that, at 15 years of follow-up, women with oestrogen-receptor-positive breast cancer who received adjuvant endocrine therapy (AET) with the drug tamoxifen for 5 years had a reduced risk of recurrence (risk reduction, 47%) and of related mortality (risk reduction, 29%). The yearly rate of death related to breast cancer also dropped by about one-third throughout the first 15 years (see Early Breast Cancer Trialists’ Collaborative Group *Lancet* **378**, 771–784; 2011). Women with this cancer type who did not receive this treatment had a 46.2% probability of recurrence of breast cancer at 15 years.

The ATLAS randomized trial showed that extending AET with tamoxifen treatment from 5 to 10 years reduced the risk of relapse (risk reduction, 30%) and of related mortality (risk reduction, 48%) after completion of therapy. The benefits of the treatment were reaffirmed by the aTTom randomized trial (G. Schiavon and I. E. Smith *Breast Cancer Res.* **16**, 206; 2014). These findings changed clinical practice. The American Society

of Clinical Oncology guidelines now recommend that women with oestrogen-receptor-positive breast cancer should consider 10 years of AET with tamoxifen.

Balkees Abderrahman, V. Craig Jordan *University of Texas MD Anderson Cancer Center, Houston, Texas, USA.*
bhabderrahman@mdanderson.org

Fragile ecosystems to test climate targets

At the 2015 climate summit in Paris, negotiators adopted 2 °C as the upper limit for global warming, with a view to limiting it to 1.5 °C. I suggest that more research is needed into ecosystems that are highly sensitive to temperature shifts and that deliver multiple ecosystem services, such as mountains and corals. Such work could help in the assessment of these targets and of the risks associated with climate-mitigation options such as bioenergy and geoengineering.

The Intergovernmental Panel on Climate Change is moving forward with its special report on the 1.5 °C warming and its Sixth Assessment Report. Each document will need to consider the future impacts of the two targets on biodiversity, ecosystems and humans — and what it would take to achieve the 1.5 °C target.

In the Himalayas, for example, projected mean increases of 1.8 °C, 2.2 °C and 3.7 °C in global mean surface temperatures for 2081–2100 (relative to 1986–2005) would lead to significantly greater loss of glaciers than if the projected increase is 1.5 °C or less (P. D. A. Kraaijenbrink *et al.* *Nature* **549**, 257–260; 2017). These glacier changes would affect biodiversity and human populations by altering species distributions, water regimes, farming and the risks of outburst floods from glacier lakes.

Ignacio Palomo *Basque Centre*

for Climate Change, Leioa, Spain.
ignacio.palomo@bc3research.org

A serious nonsense publishing proposal

The surge in open-access predatory journals is making it harder for contributors and readers to distinguish these from legitimate publications — a confusion that is fostered by the predatory-journal industry. One solution could be to deploy a variant of a well-established quality-control test.

The scientific community could submit replicate test articles several times a year to a wide array of open-access journals, suspect and non-suspect. These manuscripts would use the organization and language of legitimate science but would be readily identifiable as nonsense to someone in the field. The process should be undertaken by an independent group, perhaps under the auspices and oversight of the Directory of Open Access Journals or the US National Library of Medicine.

The results could then be made public to form the basis of a ‘journal integrity index’. This would avoid labelling journals as predatory and reduce the risk of legal retribution.

Such an objective assessment of legitimate editorial practice, which is currently almost impossible to verify, could help to eliminate the scourge of fake journals that is threatening the scientific enterprise.

Steven N. Goodman *Stanford University, California, USA.*
steve.goodman@stanford.edu

CORRECTION

The Outlook article ‘Combinations on trial’ (*Nature* **552**, S67–S69; 2017) overstated the number of immunotherapeutic agents in development at more than 2,400. In fact, there are roughly 2,000.

PLANETARY SCIENCE

Cometary spin-down

The rotation rate of a comet more than halved in two months — a much greater change than has previously been observed. This suggests that the comet is in a distinct evolutionary state and might soon reorient itself. [SEE LETTER P.186](#)

JESSICA AGARWAL

Kilometre-sized chunks of ice and dust known as cometary nuclei were left over from the formation of the Solar System¹. The vast majority of these objects orbit the Sun in one of two cometary reservoirs beyond the orbit of Neptune: the Kuiper belt and the Oort cloud. When an object from one of these reservoirs enters the inner Solar System, it becomes an active comet — its ice is transformed into gas and carries along embedded dust to form a diffuse envelope (coma) and tail. On page 186, Bodewits *et al.*² report a dramatic decrease in the rotation rate of comet 41P/Tuttle–Giacobini–Kresák (comet 41P) indicating that this object could soon enter a phase of rotational instability and reorientation that has never before been seen in a comet.

A rotating celestial body that orbits the Sun without being perturbed has a constant spin state — its rotation rate and the orientation of its axis of rotation relative to inertial space (represented approximately by the positions of stars) are fixed. But, in practice, many factors can change a body's spin state. These include the gravitational pull of other objects, collisions, asymmetric emission of thermal radiation from the body³ and, particularly in the case of comets, the recoil force from the asymmetric release of gas.

Gas that streams from a comet's surface accelerates the region of origin in the opposite direction, like a rocket engine (Fig. 1). If the direction of this acceleration does not cross the body's centre of mass, it will produce a turning effect called a torque. And if the time-averaged torques on all surface elements do not cancel each other out, they will alter the comet's spin state. Outgassing forces will also affect the body's orbit around the Sun⁴.

Moderate changes in rotation rate have been observed in several comets — in particular, those visited by spacecraft, for which high-quality data are available. For

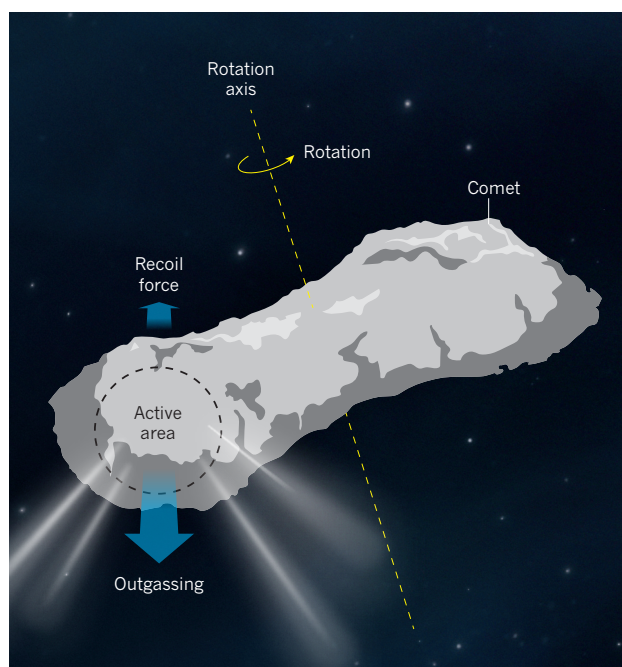


Figure 1 | Asymmetric outgassing from a comet. Bodewits *et al.*² report that the rotation rate of comet 41P/Tuttle–Giacobini–Kresák decreased rapidly between March and May 2017. They suggest that this slowdown was caused by the release of gas from a particularly active area far from the comet's rotation axis. Such asymmetric outgassing would have generated a strong recoil force, accelerating the active area in the opposite direction to the comet's rotation and thereby reducing the comet's rotation rate.

comet 67P/Churyumov–Gerasimenko, the target of the European Space Agency's Rosetta mission, a clear connection has been established between outgassing-induced torques and changes in rotation rate⁵.

If a comet is spun up to a rotation rate at which the centrifugal force near the equator surpasses gravitational and cohesive forces, landslides and partial or even catastrophic fragmentation can occur^{6–8}. Such events would be accompanied by strong sublimation (transformation of ice into gas) and dust production from newly exposed areas, which is one possible cause of sudden increases in brightness called outbursts.

Comet 41P is a small (1.4–2.0 km in diameter) body that originated from the Kuiper belt and was pulled into its current orbit in the inner Solar System by the gravity of Jupiter.

During previous passes by the Sun, known as perihelion passages, the comet had a high level of outgassing activity, given its small size⁹. It passed by Earth at only one-seventh of the Earth–Sun distance (an astronomical unit, AU) on 1 April 2017, and had its closest approach to the Sun at a distance of about 1 AU on 12 April.

Bodewits *et al.* observed comet 41P in March 2017 using the Discovery Channel Telescope at the Lowell Observatory in Arizona, and then in May using the UltraViolet–Optical Telescope on board the Swift space observatory. Over the two-month interval between their observations, the authors found that the comet's rotation period increased from an already long 20 hours to more than 46 hours. Such a high rate of change has not been seen in a comet before.

The authors conclude that comet 41P must be subject to an extremely effective torque. They suggest that this feature could be caused by outgassing from a particularly active area far from the body's rotation axis, oriented such that the gas flows in approximately the same direction as the rotation. The efficiency of the torque is enhanced by the comet's

comparatively small size, high outgassing rate and slow overall rotation.

Bodewits and colleagues extrapolated the comet's rotation period in time to explore the body's past and future spin states (see Fig. 4 of the paper²). Assuming comparable torques during past perihelion passages, the authors found that the comet could have been rotating with a period of about 5 hours, which is near the fragmentation limit, before 2006. They hypothesize that this rapid rotation might be linked to a bright outburst that occurred during the comet's 2001 perihelion passage⁹.

For instance, the rotation could have induced a landslide or partial fragmentation in the comet, which would have been visible as an outburst. Alternatively, or in addition, the event behind the outburst might have uncovered an active area that

is now causing the strong torque. A similar sequence of events could have occurred in comet 103P/Hartley 2, which was visited by the Deep Impact Extended Investigation (DIXI) space mission^{8,10} in 2010.

Extrapolating comet 41P's rotation rate forward in time, Bodewits *et al.* predict that the period would have exceeded 100 hours in mid-2017. Such an extremely slow rotation would no longer stabilize the comet's spatial orientation, so that even small torques could make it wobble like a spinning top. If the current strong torque persists, it might eventually drive the comet to spin up again, possibly about a different axis.

A change in comet 41P's rotation axis would affect the seasonal distribution of heating across the body's surface, the associated levels of activity and the pattern of mass transport between different regions¹¹. The global process of cometary erosion might therefore be redirected. Observations from the end of the 2017 activity period and from the next perihelion passage in 2022 could document this yet-to-be-seen phase of cometary evolution, and reveal valuable information about the nature of comets and other planetary bodies. ■

Jessica Agarwal is at the Max Planck Institute for Solar System Research,

37077 Göttingen, Germany.
e-mail: agarwal@mps.mpg.de

1. Whipple, F. L. *Astrophys. J.* **111**, 375–394 (1950).
2. Bodewits, D., Farnham, T. L., Kelley, M. S. P. & Knight, M. M. *Nature* **553**, 186–188 (2018).
3. Rubincam, D. P. *Icarus* **148**, 2–11 (2000).
4. Marsden, B. G. *Astron. J.* **74**, 720–734 (1969).
5. Keller, H. U., Mottola, S., Skorov, Y. & Jorda, L. *Astron. Astrophys.* **579**, L5 (2015).
6. Hirabayashi, M. *et al. Nature* **534**, 352–355 (2016).
7. Jewitt, D. *et al. Astrophys. J.* **829**, L8 (2016).
8. Steckloff, J. K., Graves, K., Hirabayashi, M., Melosh, H. J. & Richardson, J. E. *Icarus* **272**, 60–69 (2016).
9. Combi, M. *SOHO SWAN Derived Cometary Water Production Rates Collection* (NASA, 2017).
10. A'Hearn, M. F. *et al. Science* **332**, 1396–1400 (2011).
11. Keller, H. U. *et al. Mon. Not. R. Astron. Soc.* **469**, S357–S371 (2017).

NEUROSCIENCE

Neuronal plasticity in nematode worms

Neuronal activity induces changes in the connectivity of a neuron called DVB in adult male nematode worms. This discovery provides an opportunity to study a fundamental process in this powerful model organism. SEE ARTICLE P.165

SCOTT W. EMMONS

Central to the function of the nervous system is its dynamic ability to undergo changes, for instance in the physiological properties of its constituent neurons, the synaptic connections between them, and the characteristics of individual synapses. The hypothesis that neuronal activity can lead to such plasticity, first proposed by the neurophysiologist Donald Hebb in 1949, is fundamental to brain science, and has been confirmed in many studies¹. On page 165, Hart and Hobert² describe an example of experience-dependent neural plasticity in the nematode worm *Caenorhabditis elegans*, a species in which this phenomenon has been little studied³.

It is important to demonstrate this already well-described and widely studied neural phenomenon in a nematode because *C. elegans* is not just any worm, but a powerful experimental model. Genetic studies in *C. elegans* have led to the discovery of several molecular components common to all nervous systems. Furthermore, a complete map of neural connectivity in the nematode nervous system has been available for more than 30 years^{4,5} — such a connectome is not yet available for any other animal.

Assembly of the *C. elegans* connectome was made possible not only by the worm's tiny size (1 millimetre long), but also because its cells are constant in number and identity, and its synaptic connections are largely conserved

between individuals. These properties, together with the fact that connectivity data were obtained from only a few individuals, have created the impression that the *C. elegans* nervous system is exceptional in having a rigid

and constant structure. Intuition suggests that this cannot be the case — the worm's nervous system is so complex that it must be based on dynamic mechanisms. But few examples of variability in *C. elegans* neurons have been described until now.

The *C. elegans* inhibitory neuron DVB makes different connections in the worm's two sexes: males and hermaphrodites^{4,6}. A single process extends towards the head of the worm in both sexes, and a male-specific outgrowth towards the tail leads to the formation of synaptic connections to a neuron and muscles that control the movement of the male's spicules — a pair of hardened structures that insert into the vulva of the hermaphrodite during mating⁶ (Fig. 1). The formation of these new synapses, and the loss of some old ones, mean that spicule movement comes under the

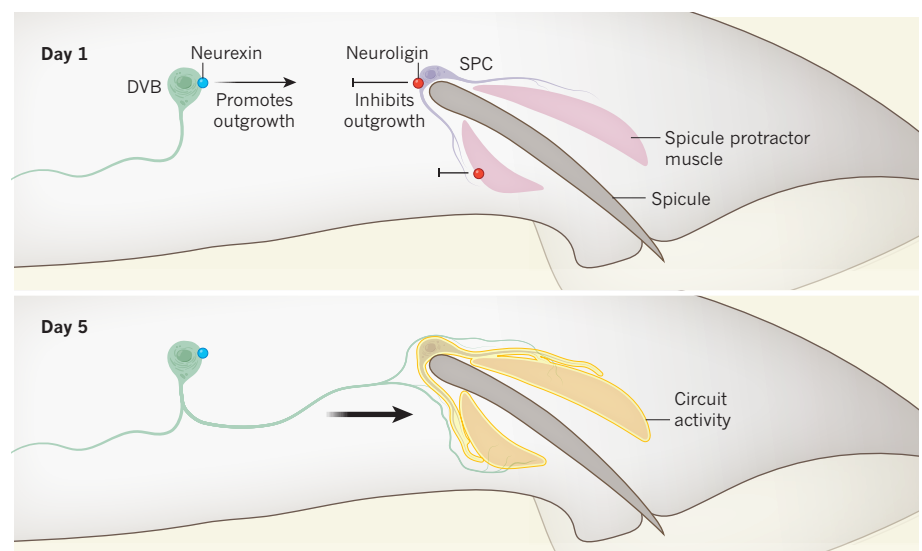


Figure 1 | Activity-dependent neuronal outgrowth in nematodes. Hart and Hobert² examined the neuron DVB in nematode worms (*Caenorhabditis elegans*). They report that, between days one and five of adulthood in male worms, DVB grows towards, and makes synaptic connections onto, spicule protractor muscles and the spicule neuron SPC, which control a male-specific mating behaviour involving movement of a structure called the spicule. This outgrowth is regulated, at least in part, by two cell-adhesion proteins: neurexin is expressed by DVB and promotes outgrowth; and neurologin is expressed by the spicule protractor muscles and SPC, and inhibits outgrowth. The authors show that the expression of neurologin is repressed when the male undergoes copulatory behaviours, activating these muscles and SPC — DVB outgrowth is therefore activity dependent.

inhibitory control of DVB. This refinement improves the male's mating efficiency⁷. Hart and Hobert now show that this male-specific outgrowth of DVB occurs between days 1 and 5 of adult life. The outgrowth produces a branching neuronal architecture that, unlike many neuronal circuits in *C. elegans*, varies between individuals.

Hart and Hobert used fluorescent 'reporter' proteins to visualize DVB outgrowth and synapse formation. Their analysis reveals that outgrowth does not occur if the male does not experience copulatory activity. The authors then mimicked natural behaviours by using sophisticated genetic techniques to activate or inhibit the signalling or movement of DVB's target neuron and muscles, respectively. This shows that activity in DVB's targets stimulates the neuron's outgrowth.

What molecular pathways might mediate DVB outgrowth? Neural cell-adhesion proteins are expressed on cell surfaces in the nervous system. They have extracellular protein-protein interaction domains that can mediate communication between cells, and are thought to have a role in encoding and building the nervous system's synaptic structure⁸. Two of the best-studied proteins in this class are neurexin and neuroligin, which can interact with one another and are involved in synapse formation and regulation⁹. As such, they were natural candidates for Hart and Hobert to test.

The authors examined the roles of these proteins by combining genetic deletion or overexpression of the proteins with stimulation or suppression of activity in the circuit. These analyses led to several findings. First, neurexin is expressed in DVB and is required for DVB outgrowth. Second, the activity of neurexin is inhibited by neuroligin, which is expressed in male sex circuits and muscles. Third, neuroligin expression is suppressed by activity in the circuit, which explains why DVB outgrowth is activity dependent. Precisely how neuroligin inhibits DVB outgrowth, and whether the two proteins physically interact in this setting, remain to be determined.

Hart and Hobert's work brings together three areas of study in neuroscience: outgrowth, branching and target selection in plastic neurons; control of these processes through neuronal activity; and the function of neural cell-adhesion proteins. The value of the study therefore lies not only in the discovery of a new phenomenon, but also in the framework it provides for making more discoveries.

Analysis of *C. elegans* mutants will make it possible to identify additional molecules that affect DVB outgrowth, such as the binding partner of neurexin that stimulates outgrowth. The intracellular mechanisms that drive DVB outgrowth, and how they are controlled by interactions between neurexin and its binding partner, can then be analysed. Other questions for study include how DVB knows where

to send processes, how its axonal extensions recognize appropriate synaptic targets, and precisely how circuit activity controls neuroligin expression.

Finally, Hart and Hobert found that these events occur only in males. The authors attempted to stimulate DVB outgrowth in hermaphrodites, but their results suggest that neither circuit activity nor the neurexin-neuroligin pathway are by themselves sufficient to do this. Other work¹⁰ in *C. elegans* suggests that it is the complement of sex chromosomes (two X chromosomes in the hermaphrodite and only one in the male) in the cells of the circuit that ultimately makes them respond to sex-neutral pathways in sex-appropriate ways.

Genetic studies⁹ have implicated mutations in neural cell-adhesion genes, including neurexin and neuroligin, as the bases of psychiatric disorders, partly because of the roles of these genes in neural plasticity. Progress in unravelling details of the molecular pathways underlying their activity could therefore have profound implications

for understanding not only learning and memory, but also mental disorders and their sex-specific expression. ■

Scott W. Emmons is in the Department of Genetics and the Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, New York, New York 10461, USA.

e-mail: scott.emmons@einstein.yu.edu

1. Holtmaat, A. & Svoboda, K. *Nature Rev. Neurosci.* **10**, 647–658 (2009).
2. Hart, M. P. & Hobert, O. *Nature* **553**, 165–170 (2018).
3. Jin, Y. & Qi, Y. B. *Curr. Opin. Neurobiol.* **48**, 97–105 (2017).
4. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. *Phil. Trans. R. Soc. B* **314**, 1–340 (1986).
5. Emmons, S. W. *Phil. Trans. R. Soc. B* **370**, 20140309 (2015).
6. Jarrell, T. A. *et al. Science* **337**, 437–444 (2012).
7. LeBoeuf, B. & Garcia, L. R. *G3 Genes Genomes Genet.* **7**, 647–662 (2017).
8. Williams, M. E., de Wit, J. & Ghosh, A. *Neuron* **68**, 9–18 (2010).
9. Südhof, T. C. *Cell* **171**, 745–769 (2017).
10. Portman, D. S. J. *Neurosci. Res.* **95**, 527–538 (2017).

ORGANOMETALLIC CHEMISTRY

Dogma-breaking catalysis

The catalysts conventionally used for industrially important hydrogenation reactions are expensive and generate toxic residues. Catalysts have now been reported that might lead to cheaper, less toxic alternatives.

DOUGLAS W. STEPHAN

Reactions of hydrogen gas with organic compounds are performed on a large scale worldwide by the chemicals industry¹. Such hydrogenation reactions are essential to the production of numerous commercial goods, including many polymers, foodstuffs and pharmaceuticals. However, a catalyst is needed to provide a thermodynamically accessible reaction pathway that allows hydrogenations to occur. Until the past decade or so, it was thought that these catalysts must derive from transition metals, but there is now a growing list of alternatives. Writing in *Nature Catalysis*, Bauer *et al.*² add to that list by reporting effective hydrogenation catalysts derived from alkaline-earth elements — the group of metals that includes calcium.

About 100 years ago, the chemist Paul Sabatier was the first to recognize that amorphous metals could act as catalysts to mediate the hydrogenation of organic substrates³. By the middle of the twentieth century, the emergence of the subdiscipline of organometallic

chemistry led to the development of a wide variety of transition-metal complexes that are highly effective catalysts for these reactions⁴. Soluble transition-metal catalysts have undergone continual development to offer higher and higher reactivities. In addition, judicious changes to the ligand molecules bound to the metal atom were found to control the reactivity and selectivity of the catalytic complexes — not only the substrate selectivity, but also the stereoselectivity (the 3D geometric arrangement of atoms generated in the product). Despite these advances, most catalysts used in industrial processes are derived from the metals platinum, palladium, rhodium and ruthenium, which are expensive, toxic and rare.

The cost of the precious metals in such catalysts is not the only expense associated with their use — the removal of toxic catalyst residues from the products is also costly. This, together with increasing environmental concerns, has prompted efforts to find alternatives to conventional hydrogenation catalysts. One strategy that uses the principles of

organometallic chemistry has been to develop catalysts derived from earth-abundant, less-toxic transition metals such as iron, cobalt and nickel⁵.

In the past decade or so, startling strategies for hydrogenation reactions have also been discovered. In 2006, certain molecules containing boron and phosphorus were shown to react reversibly with hydrogen⁶. It was subsequently found⁷ that reactions between boron-containing molecules known as boranes and phosphorus-containing molecules called phosphines can be frustrated electronically or through steric effects (which occur when bulky chemical groups block access to certain parts of a molecule). This allows certain combinations of boranes and phosphines to chemically activate hydrogen molecules, and, in some cases, mediate the hydrogenation of many different types of compound^{8,9}. Then, in 2008, a remarkable calcium-based catalyst was reported¹⁰ for the hydrogenation of alkenes (hydrocarbons that contain carbon–carbon double bonds). Collectively, these findings provided evidence that hydrogenation can be catalysed by systems based on elements other than the transition metals, overturning 100 years of chemical dogma.

Bauer *et al.* have now broadened the range of alkaline-earth-metal derivatives that can form the basis of hydrogenation catalysts. The new catalysts are complexes with the general formula $M[N(\text{SiMe}_3)_2]_2$ (where M can be magnesium, calcium, strontium or barium; Si is silicon; and Me represents a methyl group), and can be readily prepared. The authors used them to hydrogenate substrates known as aldimines (Fig. 1).

The researchers performed 30 reactions using different reaction conditions and several aldimines. They varied the amount of catalyst used (between 2.5% and 10% molar equivalents of the reaction substrate), the pressure of hydrogen (1–12 bar) and the temperature (80–120 °C). Most of the reactions were 99% complete in times ranging from 15 minutes to 24 hours, depending on the specific substrate, catalyst and conditions.

The authors show that the hydrogenations are slower for bulkier aldimine molecules and when the carbon atom in the aldimine's imine (C=N) group is less electrophilic (attractive to negative charges). Conversely, the catalytic activity increases with the atomic size of the metal used: the magnesium catalyst is least reactive, and the calcium, strontium and barium catalysts are increasingly reactive. That said, the calcium catalyst¹⁰ previously reported by researchers from the same group was a highly effective catalyst for aldimine hydrogenation, which suggests that the activity of the current calcium catalyst could be optimized by modifying the ligands bound to the metal atom.

Most of the aldimines tested with the catalysts

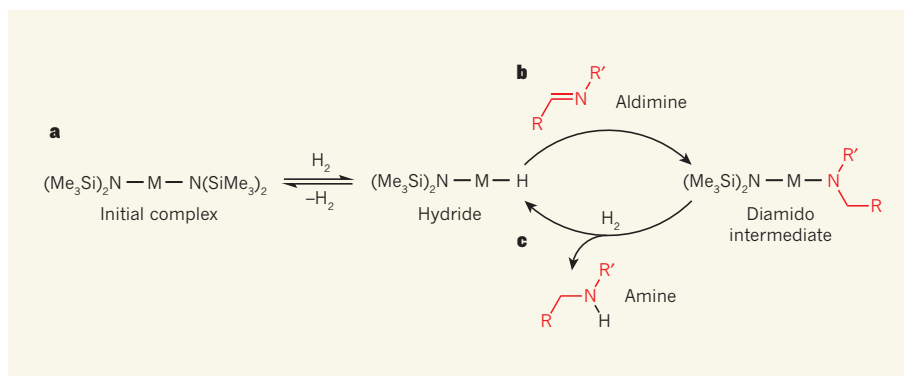


Figure 1 | Simplified hydrogenation mechanism for catalysts that contain alkaline-earth metals.

Bauer *et al.*² report that complexes containing alkaline-earth metals catalyse the reaction of aldimine compounds with hydrogen gas (H_2), and propose the following mechanism. **a**, The initial complex reacts with H_2 — probably reversibly — to generate a transient hydride intermediate. **b**, The aldimine inserts into the M–H bond of the hydride to form a diamido intermediate. **c**, This intermediate reacts with more H_2 to liberate the hydrogenation product (an amine), regenerating the hydride for further catalytic cycles. Me, methyl; Si, silicon; M, magnesium, calcium, strontium or barium; R is typically an aromatic group, such as phenyl; R' is *t*-butyl, isopropyl, phenyl or a mesityl (a bulky analogue of the phenyl group).

had a phenyl group (a benzene ring) attached to the carbon atom in the imine. Bauer *et al.* found that the catalysts still worked when electron-withdrawing or electron-donating groups were attached to the phenyl group — something that isn't always guaranteed in chemical reactions. However, the catalysts could not hydrogenate compounds known as ketimines, which are similar to aldimines but have two groups attached to the imine carbon atom, rather than just one.

Bauer and co-workers propose a mechanism for the catalytic cycle in which the catalyst first

reacts with hydrogen gas to generate a transient hydride intermediate — a process that is likely to be reversible (Fig. 1a). The aldimine inserts into the M–H bond of the hydride to generate a diamido intermediate (Fig. 1b),

which then reacts with more hydrogen gas to liberate the hydrogenation product (an amine; Fig. 1c). This last step also regenerates the hydride for further catalytic cycles.

The proposed mechanism might seem straightforward, but the authors note that the active form of the catalyst has not been unambiguously identified. When Bauer *et al.* reacted the calcium catalyst with hydrogen gas alone, the proposed hydride intermediate did form, but so, too, did aggregated forms of hydrides.

The researchers performed computational simulations of their reactions to cast further light on the reaction mechanism. The simulations revealed that the aggregation process probably releases energy, suggesting that a thermodynamic driving force could generate

a currently undefined, catalytically active complex involving an aggregated hydride. The simulations also supported the proposed stepwise mechanism for the catalytic cycle. Finally, Bauer *et al.* used a previously reported calcium hydride complex¹¹ as a model of the proposed catalytic hydride intermediate, and found that it reacts with an aldimine and hydrogen gas in a way that is consistent with the proposed catalytic cycle.

Compared with industrial reactions catalysed by transition-metal complexes, Bauer and colleagues' reactions use higher amounts of catalyst and are relatively slow. Nonetheless, the findings expand the substrate scope for hydrogenation catalysts derived from abundant alkaline-earth metals, raising the possibility that low-cost and low-toxicity catalysts could one day be used for industrial applications. Work is now needed to improve the activity of such catalysts, and to find catalysts that tolerate the presence of the impurities found in industrial-grade reagents. Perhaps most crucially, the authors' work provides further evidence that the dogma that transition metals are required for hydrogenation catalysis should be firmly relegated to the false beliefs of the past. ■

Douglas W. Stephan is in the Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6 Canada.
e-mail: douglas.stephan@utoronto.ca

1. de Vries, J. G. & Elsevier, C. J. (eds) *Handbook of Homogeneous Hydrogenation* (Wiley, 2006).
2. Bauer, H. *et al.* *Nature Catal.* <http://dx.doi.org/10.1038/s41929-017-0006-0> (2018).
3. Sabatier, P. *Ind. Eng. Chem.* **18**, 1005–1008 (1926).
4. Hudlický, M. *Reductions in Organic Chemistry* 2nd edn (Am. Chem. Soc., 1996).
5. Bullock, R. M. (ed.) *Catalysis without Precious Metals*

- (Wiley, 2010).
 6. Welch, G. C., San Juan, R. R., Masuda, J. D. & Stephan, D. W. *Science* **314**, 1124–1126 (2006).
 7. Welch, G. C. & Stephan, D. W. *J. Am. Chem. Soc.* **129**, 1880–1881 (2007).

8. Chase, P. A., Welch, G. C., Jurca, T. & Stephan, D. W. *Angew. Chem. Int. Edn* **46**, 8050–8053 (2007).
 9. Stephan, D. W. & Erker, G. *Angew. Chem. Int. Edn* **54**, 6400–6441 (2015).
 10. Spielmann, J., Buch, F. & Harder, S. *Angew. Chem.*

- Int. Edn* **47**, 9434–9438 (2008).
 11. Causero, A. *et al. Organometallics* **35**, 3350–3360 (2016).

This article was published online on 8 January 2018.

BIOTECHNOLOGY

An ode to gene edits that prevent deafness

Gene editing can prevent inherited deafness in mice by disabling a mutant version of a gene that causes hearing loss. Is this a turning point on the path towards treating some types of human deafness? [SEE LETTER P.217](#)

FYODOR URNOV

When the 32-year-old composer Ludwig van Beethoven realized that his hearing was failing, he wrote to his brothers that “as the leaves of autumn wither and fall, so has my own life become barren”. Although the cause of Beethoven’s deafness is unknown, there are many examples of hearing loss in later life that are linked to inherited DNA changes. Two centuries later, techniques to prevent inherited forms of deafness are finally getting closer to implementation in the clinic. On page 217, Gao *et al.*¹ report progress in using gene-editing technology to treat a mouse model of inherited deafness. Given the growing momentum in using genetic engineering for human therapy, the path needed to take this approach to the clinic is clear.

The remarkable process of sensing sound occurs in the inner ear². Tiny, hair-like structures called cilia on the surface of hair cells in the cochlea respond to sound waves. Ciliary motion evokes an electrical signal because the properties of a protein assembly at the base of each cilium change when such motion occurs. The *Tmc1* protein is thought³ to be part of this assembly in humans, and some *TMC1* mutations cause people to lose their hearing over time. The symptoms start in childhood, and deafness, along with associated degeneration and death of hair cells, ensues within 10 to 15 years⁴.

Gao and colleagues analysed the Beethoven mouse strain, in which the animals have a *Tmc1* mutation that causes them to grow deaf over time⁵. The mouse mutation they studied matches a mutation in human *TMC1* that is also linked to progressive hearing loss⁶. The mutation is dominant, which means that even if only one of a person’s two copies of the gene has the mutation, they will become deaf. The mutant copy of the gene produces a defective protein that somehow impairs cell function, even though the cell also

has a wild-type copy of the gene⁷.

The repair of dominant-mutation-associated deafness is a delicate matter — the mutated gene must be disabled while preserving the wild-type gene within the same cell. This is no trivial undertaking, because only one nucleotide of DNA distinguishes the two

versions of the *TMC1* gene from each other (Fig. 1). One way to understand this is to imagine a duet between two people trying to sing in unison. If one person is off-key, this offender must be selectively silenced to allow the correct tune to be heard, because if both singers are stopped, the music will cease.

Gene editing is the technique of choice to rid a hair cell of the mutant version of a gene⁸. This involves using a nuclease enzyme to cut a targeted DNA sequence in the specific gene inside the living cell. The cut causes a double-strand DNA break and the repair process often results in mistakes in which nucleotides are added or lost. Such a change can alter the sequence in a way that might cause translation to prematurely arrest and thereby prevent gene expression.

The authors used the nuclease Cas9, which cuts DNA at a specific site by using a snippet

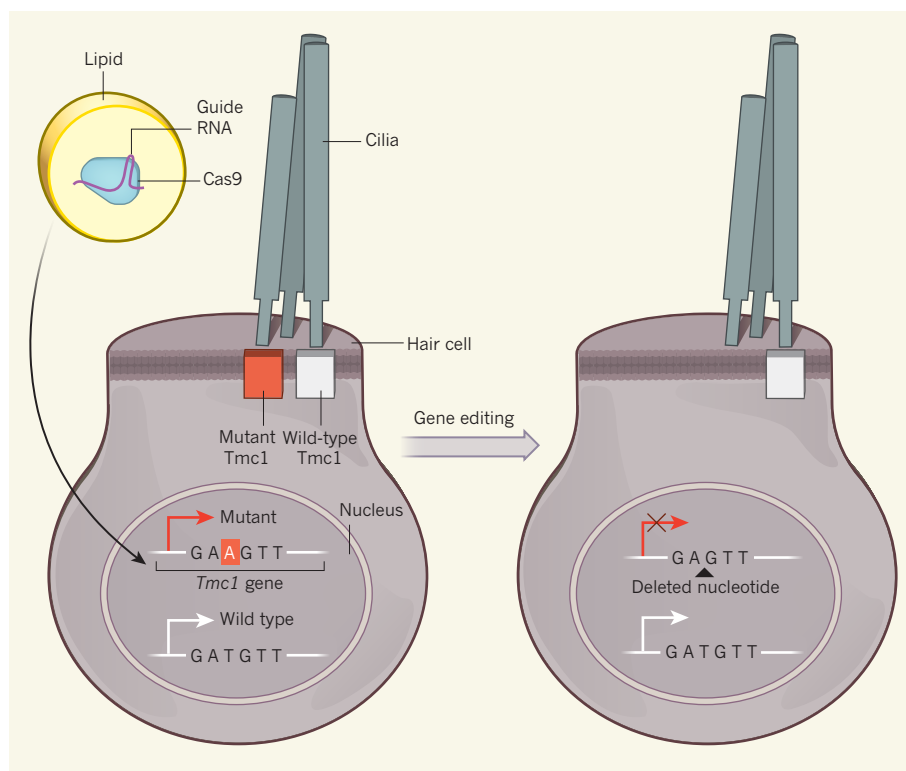


Figure 1 | Gene editing in mice can prevent inherited hearing loss. Gao *et al.*¹ investigated a mouse model of later-life deafness that is caused by a mutant version of the *Tmc1* gene. This mutation is identical to one in the human version of the gene that is linked to deafness. Hearing loss is accompanied by the death of inner-ear hair cells that sense sound using their ciliary projections. The authors injected the ears of newborn mice with gene-editing components: the nuclease enzyme Cas9 that can cut DNA, and a guide RNA that targets Cas9 to the mutant version of *Tmc1* in hair-cell nuclei. These were packaged in a lipid droplet that fuses with cells to enable the gene-editing components to enter. The mutant version of *Tmc1* has an adenine nucleotide (A, highlighted in red in the mutant nucleotide sequence) at a position that is a thymine nucleotide (T) in the wild-type version. Gene editing selectively inactivated the mutant version of the gene through mechanisms such as nucleotide deletion. Edited cells express only the wild-type *Tmc1* protein (white) and don’t express the mutant version (red).

of RNA that binds to both the enzyme and the target DNA⁹. This approach is also known as CRISPR–Cas gene editing. The guide RNA matches the mutant but not the wild-type gene, enabling Gao and colleagues to solve the problem of ensuring that the mutant form of the gene is cut whereas the wild-type version is left untouched.

Another challenge was to get Cas9 into the inner ear. *In vivo* gene-editing approaches often rely on viruses to introduce nuclease-encoding sequences into the organism being edited^{10,11}. However, Gao and colleagues reasoned that, when the nuclease has done its job in the cell, it will no longer be required, so introducing the protein itself should suffice. They turned to a technique they had used previously¹², in which they packaged Cas9 protein bound to its guide RNA in a type of lipid droplet that can fuse with cells, enabling the editing machinery to enter. The authors injected these droplets into the inner ear of newborn Beethoven mice.

The inner ears of unedited adult Beethoven mice were barren of hair cells; however, their gene-edited adult siblings had inner-ear hair cells that were almost indistinguishable in shape and number from those in wild-type mice. The edited animals could be startled by a sudden loud noise, whereas their unedited siblings could not. More-sophisticated measurements also confirmed that hearing improved as a result of gene editing. Encouragingly, the engineered nuclease seems to have stayed true to its design and did not create undesired genetic changes of concern in the DNA of the hair cells.

A modest fraction of cells were edited. The authors propose that this low proportion of edited cells resulted in a beneficial ‘halo’-like effect on neighbouring unedited cells that still contained the mutant form of the gene, preventing the death and degeneration of these neighbouring cells. Although the mechanism underlying this proposed halo effect is unclear, the finding offers encouragement for the clinical adoption of this approach, because it suggests that the genetic repair of all hair cells is perhaps not needed to achieve a beneficial effect on hearing.

Gao and colleagues’ work provides an essential first step towards moving this type of approach nearer to the clinic by providing evidence that it is safe and effective in an animal that has a similar genetic mutation and comparable hearing loss to those in humans. How long could it be before individuals with this *TMC1* mutation might be treated using gene editing? One reason for optimism comes from the pace at which other gene-editing approaches have reached the clinic.

To give just a few examples from clinical trials, the gene *CCR5* has been inactivated in immune-system cells using a type of enzyme called a zinc-finger nuclease to try to reduce the viral load in people infected with

HIV¹³. Immune cells have also been edited to generate cancer-targeting cells¹⁴. However, these techniques required cells to be removed from the patient’s body for gene editing and then replaced. Ear cells cannot be removed, so a direct *in vivo* approach is needed, which is even more challenging to achieve than *ex vivo* gene editing.

Encouragingly, such *in vivo* gene editing (for a different condition) has been performed in a clinical trial using zinc-finger nucleases¹⁵, and the work leading up to that¹⁶ makes clear the next steps for Gao and colleagues’ approach. A nuclease must be found that has clinical-grade potency and specificity in human cells. Lipids must be identified that can be safely injected along with the nuclease into the human inner ear. Next, this nuclease must be tested for safety in larger animals, such as primates. An *in vivo* virus-based gene therapy for direct injection into the eye¹⁷ has been recommended for approval in the United States, and that work provides a road map for the scientific, medical and commercial considerations that need to be taken into account when moving to the clinic.

In 1902, the physician Archibald Garrod initiated the first study that demonstrated a link between a gene and a disease. Since then, more than 5,000 diseases have been linked to single-gene changes. However, without the tools to modify disease-causing forms of genes, geneticists have often been unable to see their knowledge put to use for clinical benefit. The

progress being made with genome editing is changing this. Although Beethoven never heard his famous *Ode to Joy*, it could be that — thanks in no small part to his murine namesake’s fateful encounter with Cas9 — we are getting closer to the day when individuals with deafness-causing mutations can be treated by gene editing to prevent hearing loss. ■

Fyodor Urnov is at the *Altius Institute for Biomedical Sciences, Seattle, Washington 98121, USA.*

e-mail: urnov@altius.org

1. Gao, X. *et al. Nature* **553**, 217–221 (2018).
2. Fettiplace, R. & Kim, K. X. *Physiol. Rev.* **94**, 951–986 (2014).
3. Kawashima, Y. *et al. J. Clin. Invest.* **121**, 4796–4809 (2011).
4. Kurima, K. *et al. Nature Genet.* **30**, 277–284 (2002).
5. Zhao, Y. *et al. PLoS ONE* **9**, e97064 (2014).
6. Vreugde, S. *et al. Nature Genet.* **30**, 257–258 (2002).
7. Pan, B. *et al. Neuron* **79**, 504–515 (2013).
8. Carroll, D. *et al. Annu. Rev. Biochem.* **83**, 409–439 (2014).
9. Jiang, F. & Doudna, J. A. *Annu. Rev. Biophys.* **46**, 505–529 (2017).
10. Li, H. *et al. Nature* **475**, 217–221 (2011).
11. Ran, F. A. *et al. Nature* **520**, 186–191 (2015).
12. Zuris, J. A. *et al. Nature Biotechnol.* **33**, 73–80 (2014).
13. Tebas, P. *et al. N. Engl. J. Med.* **370**, 901–910 (2014).
14. Qasim, W. *et al. Sci. Transl. Med.* **9**, eaaj2013 (2017).
15. <https://clinicaltrials.gov/ct2/show/NCT03041324>
16. Sharma, R. *et al. Blood* **126**, 1777–1784 (2015).
17. Russell, S. *et al. Lancet* **390**, 849–860 (2017).

This article was published online on 20 December 2017.

MATERIALS SCIENCE

Rule-breaking perovskites

A material from the perovskite family of semiconductors emits light much more efficiently than expected. The explanation for this anomalous behaviour could lead to improvements in light-emitting technology. SEE LETTER P.189

MICHELE SABA

When a semiconductor absorbs light, a particle-like entity called an exciton can be produced. Excitons comprise an electron and a hole (the absence of an electron), and have two possible states: singlet and triplet. Triplet states were thought to be poor emitters of light, but, on page 189, Becker *et al.*¹ report that semiconductors known as lead halide perovskites have bright triplet excitons. The results could signify a breakthrough in optoelectronics because triplet states are three times more abundant than singlet states² and currently limit the efficiency of organic light-emitting diodes³.

Conventional wisdom holds that triplet states are dark because of the spin selection rule⁴, which forbids electrons from changing their intrinsic angular momentum (spin) during an optical transition — the process in which an atom or molecule switches from one energy state to another by emitting or absorbing light. The rule is taught in quantum-mechanics classes when atomic transitions are first introduced, and is so general that one might think that it is written in stone. Fortunately, there are loopholes that can be exploited.

The search for emissive triplet states has focused on a certain principle of quantum mechanics: if an electron’s spin is coupled

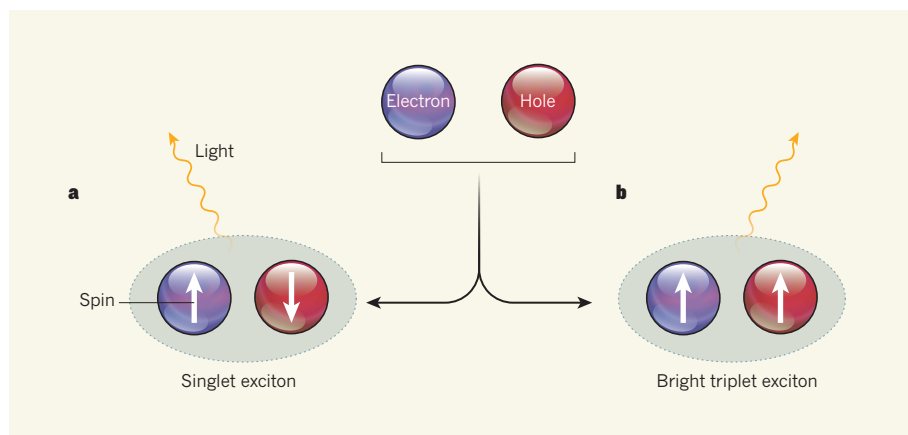


Figure 1 | Exciton emission. Semiconductors contain electrons and holes (absences of electrons) that can combine to form bound states called excitons. **a**, In singlet excitons, the intrinsic angular momenta (spins) of the electron and the hole point in opposite directions, which facilitates the emission of light. **b**, Conversely, in triplet excitons, the spins point in the same direction. Conventional wisdom holds that such states are dark, but Becker *et al.*¹ report that semiconductors known as lead halide perovskites emit light through bright triplet excitons.

to another form of angular momentum (namely, orbital momentum), the sum of the two momenta needs to be conserved, rather than the spin alone. The effect is known as spin–orbit coupling in atomic physics and as intersystem crossing in the study of organic semiconductors. It is responsible for weak emission from triplet states in atoms and organic molecules, especially when heavy elements are involved. However, until now, the strength of triplet emission was thought always to be inferior to that of singlet emission.

Lead halide perovskites seem to dispose of all conventional wisdom in materials science. Like organic semiconductors, they are relatively easy to fabricate, and their bandgap (a property that determines their conductivity and optical properties) can be tuned by varying their composition. Yet, like thin-layer (epitaxial) inorganic semiconductors, they are highly crystalline and exhibit efficient charge transport. It is as if their properties were selected from a materials scientist's wish list, combining the best aspects of organic molecules, nanocrystals and epitaxial inorganic semiconductors.

Becker and colleagues' study suggests that there is another feature of lead halide perovskites to be added to this list. The authors used a combination of theoretical and experimental work to show that nanocrystals of caesium lead halide perovskites (CsPbX_3 , where X is chlorine, bromine or iodine) have bright triplet excitons (Fig. 1). This property results in an emission rate surpassing that of other known nanocrystals⁵.

The energy difference between the triplet and singlet states in CsPbX_3 nanocrystals is relatively small (of the order of 1 millielectronvolt). Becker *et al.* therefore explored the material's emission at cryogenic temperatures (a few kelvin), to prevent

transitions between triplet and singlet states. It is unclear to what extent bright triplet states affect the material's emission efficiency at room temperature — when thermal energy greatly exceeds the singlet–triplet splitting energy and all states are equally populated. Nevertheless, the authors' findings are of fundamental relevance.

Future work will certainly investigate whether bright triplet states exist in other types of perovskite, such as hybrid perovskites that have organic, positively charged ions (cations). Such materials include the archetypal methylammonium lead iodide ($\text{CH}_3\text{NH}_3\text{PbI}_3$), and are typically prepared not as nanocrystals,

“The authors’ study highlights the potential of perovskite materials as efficient light emitters.”

but as solid-state films⁶. Unlike CsPbX_3 nanocrystals, these films comprise micrometre- or millimetre-sized crystalline domains, in which excitons dissociate into pairs of free electrons and holes at room temperature. More generally, Becker and colleagues' theoretical analysis might help scientists to identify other semiconducting materials (either organic or inorganic) that have bright triplet excitons.

Research into hybrid perovskites has been fuelled in the past few years by the successful incorporation of these materials into solar cells. Such devices can now convert more than 22% of the energy received from sunlight into electricity⁷, which is a record for perovskite solar cells. However, because of a concept known as quantum-mechanical reciprocity, there is an unavoidable energy loss in solar cells: that due to photoluminescence, which is the reverse of the absorption process⁸. As a consequence, the best solar cells are also the

best light emitters — an idea reinforced by Becker and colleagues' work.

Perovskite solar cells are now leaving academic labs and entering the market, thanks to substantial industrial efforts. The competition is mainly silicon solar cells, which have become so cheap that they negate some of the advantages of perovskite fabrication. For this reason, tandem solar cells (consisting of two sub-cells) and innovative architectures involving perovskites are being developed that can outperform commercial silicon devices in terms of efficiency, if not cost⁹.

Light emission is an application in which organic semiconductors and nanocrystals have already found commercial success, because of their ability to produce vivid colours and to be incorporated into thin panels. And yet the electric-current densities in organic light-emitting diodes are much lower than in their inorganic counterparts as a result of poor electrical conductivity. Perovskites could allow high current densities and efficiencies to be realized on large-area, thin panels¹⁰.

Becker and colleagues' study highlights the potential of perovskite materials as efficient light emitters. Although the findings might seem surprising at first sight, they should be seen as a natural consequence of quantum-mechanical reciprocity — that the class of material brought to the forefront by solar-cell technology could find applications in light emission. ■

Michele Saba is in the Department of Physics, University of Cagliari, I-09042 Monserrato, Italy.

e-mail: saba@unica.it

1. Becker, M. A. *et al.* *Nature* **553**, 189–193 (2018).
2. Pope, M. & Swenberg, C. E. *Electronic Processes in Organic Crystals* (Oxford Univ. Press, 1999).
3. Reineke, S., Thomschke, M., Lüssem, B. & Leo, K. *Rev. Mod. Phys.* **85**, 1245–1293 (2013).
4. Demtröder, W. *Atoms, Molecules and Photons* (Springer, 2010).
5. Rainò, G. *et al.* *ACS Nano* **10**, 2485–2490 (2016).
6. Stoumpos, C. C. & Kanatzidis, M. G. *Acc. Chem. Res.* **48**, 2791–2802 (2015).
7. Correa-Baena, J.-P. *et al.* *Science* **358**, 739–744 (2017).
8. Rau, U. *Phys. Rev. B* **76**, 085303 (2007).
9. Eperon, G. E., Hörantner, M. T. & Snaith, H. J. *Nature Rev. Chem.* **1**, 0095 (2017).
10. Colella, S., Mazzeo, M., Rizzo, A., Gigli, G. & Listorti, A. *J. Phys. Chem. Lett.* **7**, 4322–4334 (2016).

CORRECTION

The News & Views article ‘Cancer: Tumour lymph vessels boost immunotherapy’ by Christine Moussion and Shannon J. Turley (*Nature* **552**, 340–342; 2017) cited reference 2 incorrectly. The correct reference is: S. L. Topalian, C. G. Drake & D. M. Pardoll *Cancer Cell* **27**, 450–461 (2015).

Neurexin controls plasticity of a mature, sexually dimorphic neuron

Michael P. Hart¹ & Oliver Hobert¹

During development and adulthood, brain plasticity is evident at several levels, from synaptic structure and function to the outgrowth of dendrites and axons. Whether and how sex impinges on neuronal plasticity is poorly understood. Here we show that the sex-shared GABA (γ -aminobutyric acid)-releasing DVB neuron in *Caenorhabditis elegans* displays experience-dependent and sexually dimorphic morphological plasticity, characterized by the stochastic and dynamic addition of multiple neurites in adult males. These added neurites enable synaptic rewiring of the DVB neuron and instruct a functional switch of the neuron that directly modifies a step of male mating behaviour. Both DVB neuron function and male mating behaviour can be altered by experience and by manipulation of postsynaptic activity. The outgrowth of DVB neurites is promoted by presynaptic neurexin and antagonized by postsynaptic neuroligin, revealing a non-conventional activity and mode of interaction of these conserved, human-disease-relevant factors.

Experience modifies the structure and function of neurons and circuits in the brain through multiple mechanisms of neuronal plasticity^{1,2}. Plasticity in adult brains refines circuits in response to experience in order to mediate adaptation and homeostasis, and as a cellular correlate of learning and memory^{1,3,4}; this type of plasticity includes extension and retraction of dendrites and axons^{5–7}. The molecular mechanisms that underlie morphological plasticity in adult neurons are not well understood. Similarly, though the sexual identity of an organism influences the function and plasticity of its nervous system, the molecular and cellular bases of such sexual dimorphism are also not fully understood.

Morphological plasticity in adult male DVB neuron

The GABAergic motor neuron/interneuron DVB is located in the tail of *C. elegans* and projects anteriorly in the ventral nerve cord in both sexes (Fig. 1a). We used fluorescent reporter gene technology to visualize DVB and found that it displays extensive post-developmental morphologic plasticity exclusively in males, characterized by the progressive extension of new neurites posteriorly into the tail (Fig. 1b; Extended Data Fig. 1). The total neurite length and the number of neurite junctions increase significantly ($P < 0.001$) from day 1 to day 5 of adult life (Fig. 1c, d). The branching pattern of male DVB neurites lacks any overt stereotypy (Extended Data Fig. 2a, b). The generation of new DVB neurites in males is accompanied by the addition of presynaptic boutons containing the synaptic marker RAB-3, suggesting that these neurites are axon-like (Fig. 1b, Extended Data Fig. 1); electron microscopy analysis supports this conclusion^{8,9}. We have not identified other neurons that undergo comparable neurite outgrowth in adulthood (Fig. 1b, Extended Data Fig. 2c–h).

Dimorphic DVB connectivity influences behaviour

In hermaphrodite worms, DVB controls defecation behaviour¹⁰; in males it also contributes to protraction of the male-specific spicule structures, which are inserted into the hermaphrodite vulva during copulation¹¹ (Fig. 1e–g). Consistent with a sexually dimorphic function, the synaptic wiring pattern of DVB is also notably sexually dimorphic^{8,9} (Fig. 1g). To test for functional roles of DVB neurite outgrowth, we examined DVB function over the period of DVB neurite outgrowth.

Day 1 males have been shown to protract their spicules briefly following the expulsion step of defecation, owing to connections between defecation and spicule circuits¹¹. This seemingly pointless protraction can result in chronic protraction of spicules, which is detrimental to male mating ability. We found that day 1 males, but not day 3 males, frequently protracted spicules during expulsion¹² (Extended Data Fig. 3b). To determine whether DVB was involved in this change, we silenced DVB using expression of a histamine-gated chloride channel (*lim-6^{int4}::HisCl1* with histamine), which resulted in increased protraction of spicules with expulsion at day 3 (Extended Data Fig. 3b). The time between consecutive expulsions was unchanged between day 1 and day 3 in controls, but slightly increased in DVB-silenced day 3 males (Extended Data Fig. 3c). These results suggest that DVB plays a role in reducing expulsion-associated spicule protraction during the period of neurite outgrowth, probably through inhibition of spicule circuit components that connect with the defecation circuit. Moreover, laser ablation of DVB in day 1 males (Extended Data Fig. 3d) resulted in a reduction in the number of males with chronically protracted spicules compared to controls, whereas ablation of DVB on each day after day 2 resulted in a progressive increase in worms with chronically protracted spicules (Fig. 2a). Thus, DVB contributes to spicule protraction at day 1 and inhibits spicule protraction after day 2, with a functional consequence of suppressing spicule protraction during expulsion.

We validated these findings using expression of channelrhodopsin in DVB (Extended Data Fig. 3e). Light-induced activation of DVB in day 1 adult males resulted in observable movement of spicules, whereas activation of DVB at day 5 resulted in only rare movement of spicules (Fig. 2b, Supplementary Video 1). Expression and activation of channelrhodopsin in the spicule protraction neurons and muscles always resulted in spicule protraction at days 1 and 5 (Fig. 2b, Supplementary Video 2, Extended Data Fig. 3f). The fraction of male worms exhibiting spicule movement after channelrhodopsin-mediated DVB activation at day 1 was unchanged in males lacking GABA signalling components (*unc-25/GAD* or *unc-49/GABA_A* receptor mutants; Fig. 2c), indicating that DVB may signal through electrical connections and/or neuropeptides¹¹. Although DVB neurite outgrowth was not affected in *unc-49* mutants, these worms did show a reduction in

¹Department of Biological Sciences, Howard Hughes Medical Institute, Columbia University, New York, New York, USA.

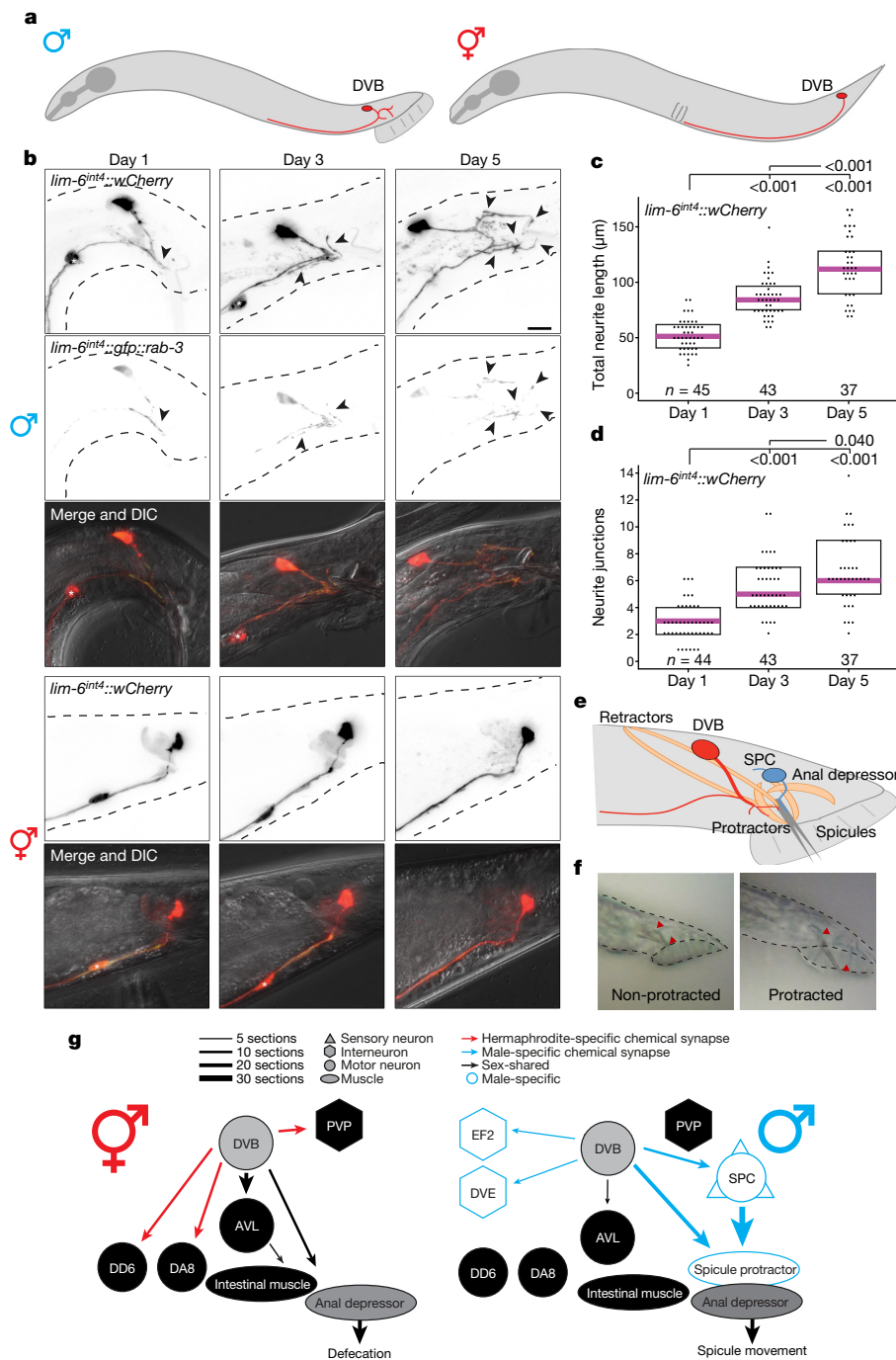


Figure 1 | Progressive neurite outgrowth of the GABAergic DVB neuron in adult males.

a, DVB neuron schematic. **b**, DVB visualized with *lim-6^{int4}::wCherry* in adult males and hermaphrodites (asterisk, PVT neuron; arrowheads, DVB neurites, n as indicated in **c**). Presynaptic boutons visualized with presynaptic marker *lim-6^{int4}::gfp::rab-3*. DIC, differential interference contrast. **c**, **d**, DVB neurite outgrowth in males quantified by total neurite length (**c**) and number of neurite junctions (**d**). Dot represents one worm; magenta bar, median; boxes, quartiles. Comparison using one-way ANOVA and post-hoc Tukey HSD, P values shown above plots, bold shows significance ($P < 0.05$). **e**, Schematic of DVB and postsynaptic spicule-associated neurons and muscles in male tail. **f**, Sample images of males with non-protracted or protracted spicules (red triangles indicate base and tip of spicules; tail and male fan are outlined, demonstrating protracted spicules extending underneath male fan; $n > 10$). **g**, Connectivity of DVB at adult stage inferred from electron micrographs (sections indicate number of EM sections over which *en passant* synapses were observed)^{8,11}. Behavioural output indicated for each sex. Scale bars, 10 μ m.

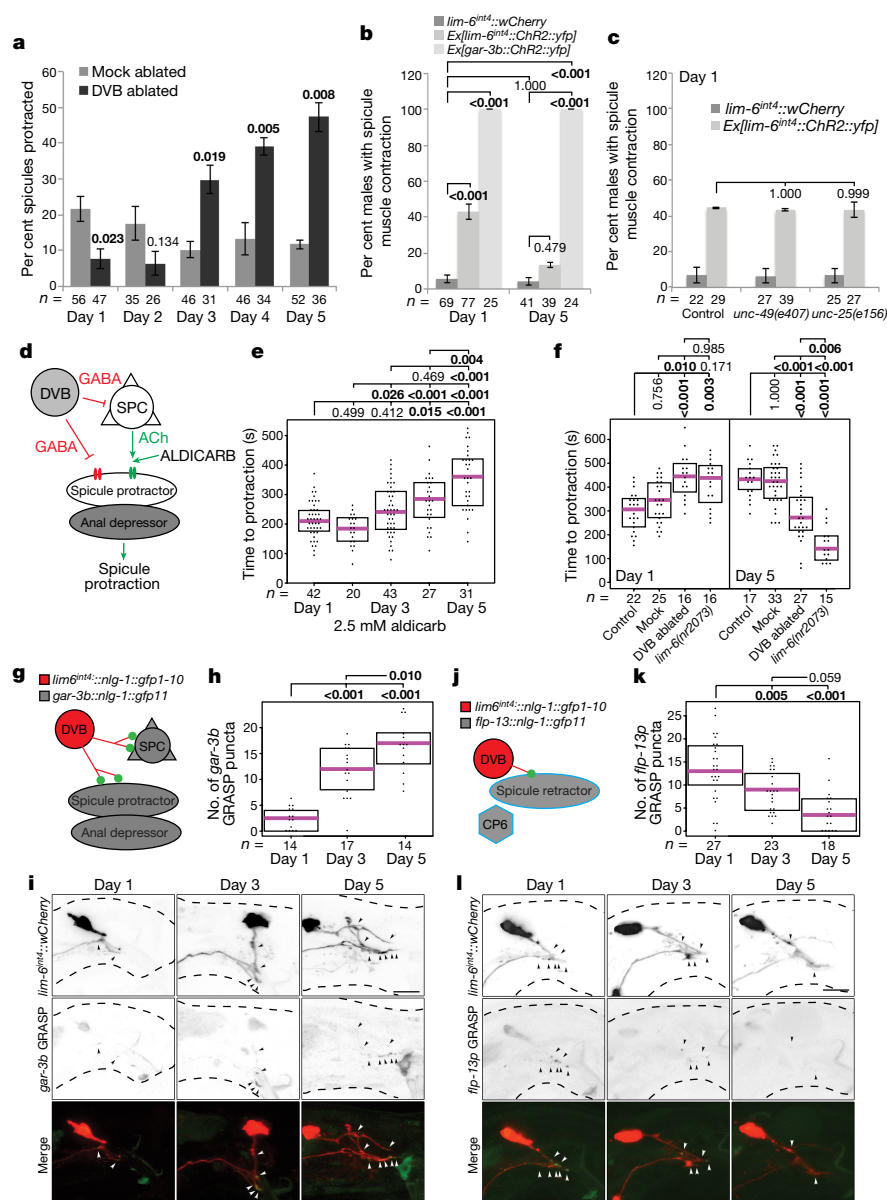
spicule protraction at day 5 (aldicarb assay described below, Extended Data Fig. 4a–d), suggesting that GABA contributes to restriction of spicule protraction in later adulthood.

To further characterize the role of DVB in active spicule protraction, we used the acetylcholine esterase inhibitor aldicarb, which induces spicule protraction through the accumulation of acetylcholine at neuromuscular synapses onto spicule protractor muscles¹³ (Fig. 2d). Aldicarb-induced spicule protraction took longer as males aged from day 1 to day 5 (Fig. 2e), during the same period as DVB neurite outgrowth. To directly test whether DVB is involved in this behavioural change, we combined laser ablation of DVB with aldicarb-induced spicule protraction. DVB ablation at day 1 resulted in slower spicule protraction in response to aldicarb than in control and mock-ablated males, again demonstrating that DVB input at day 1 has an excitatory effect on spicule protraction (Fig. 2f). DVB ablation at day 5 resulted in faster spicule protraction in response to aldicarb than in

control and mock-ablated males, demonstrating a functional switch for DVB from an excitatory to an inhibitory input on spicule protraction (Fig. 2f). These results were confirmed using genetic ‘ablation’ of DVB (*lim-6* transcription factor mutant¹⁴; Fig. 2f). Together, our results confirm that DVB switches function in adulthood, and implicate DVB as the main contributor to the temporal change observed in spicule protraction and defecation behaviour.

To investigate how the switch of DVB function during DVB neurite outgrowth relates to changes in synaptic connectivity, we used trans-synaptic labelling (GRASP¹⁵) to visualize synapses between DVB and the spicule protraction neurons and muscles (Fig. 2g). The number of these specific synaptic connections increased from day 1 to 5 (Fig. 2h, i). We also visualized synapses between DVB and the spicule retractor muscles (Fig. 2j; Extended Data Fig. 4h); the number of these synapses decreased from day 1 to 5 (Fig. 2k, l). These results provide evidence that structural remodelling of axons and dendrites in

Figure 2 | DVB neuron undergoes a functional switch in adulthood resulting in dynamic behavioural output. **a**, Per cent of mock-ablated or DVB-ablated males with chronically protracted spicules 20 h after ablation at day indicated (mean \pm s.e.m., two-tailed Student's *t*-test). **b**, Per cent of worms responding to 488-nm light with movement of spicules for control, *Ex[lim-6^{int4}::ChR2::yfp]*, and *Ex[gar-3b::ChR2::yfp]* worms (mean \pm s.e.m., one-way ANOVA and post-hoc Tukey HSD). **c**, Per cent of worms with or without *Ex[lim-6^{int4}::ChR2::yfp]* responding to blue light with spicule movement at day 1 in control, *unc-49(e407)*, and *unc-25(e156)* males (mean \pm s.e.m., one-way ANOVA and post-hoc Tukey HSD). **d**, Diagram of GABA and acetylcholine input onto spicule muscles showing site of aldicarb action. **e**, Males on 5 mM aldicarb medium timed for spicule protraction for more than 5 s. Dot represents one worm; magenta bar, median; boxes, quartiles; one-way ANOVA and post-hoc Tukey HSD. **f**, Control, mock-ablated, DVB-ablated, or *lim-6(nr2073)* mutant males timed for aldicarb-induced spicule protraction 12 h after ablation. **g**, Diagram of synaptic connections labelled with *lim-6/gar-3* GRASP. **h**, **i**, Quantification (**h**) and confocal images (**i**) of *lim-6/gar-3* GRASP puncta. **j**, Diagram of synaptic connections labelled with *lim-6/flp-13* GRASP. **k**, **l**, Quantification (**k**) and confocal images (**l**) of *lim-6/flp-13* GRASP puncta. For **i**, **l**, arrowheads, location of GRASP puncta on DVB neurites; scale bars, 10 μ m. *P* values shown above plots, bold shows significance (*P* < 0.05).



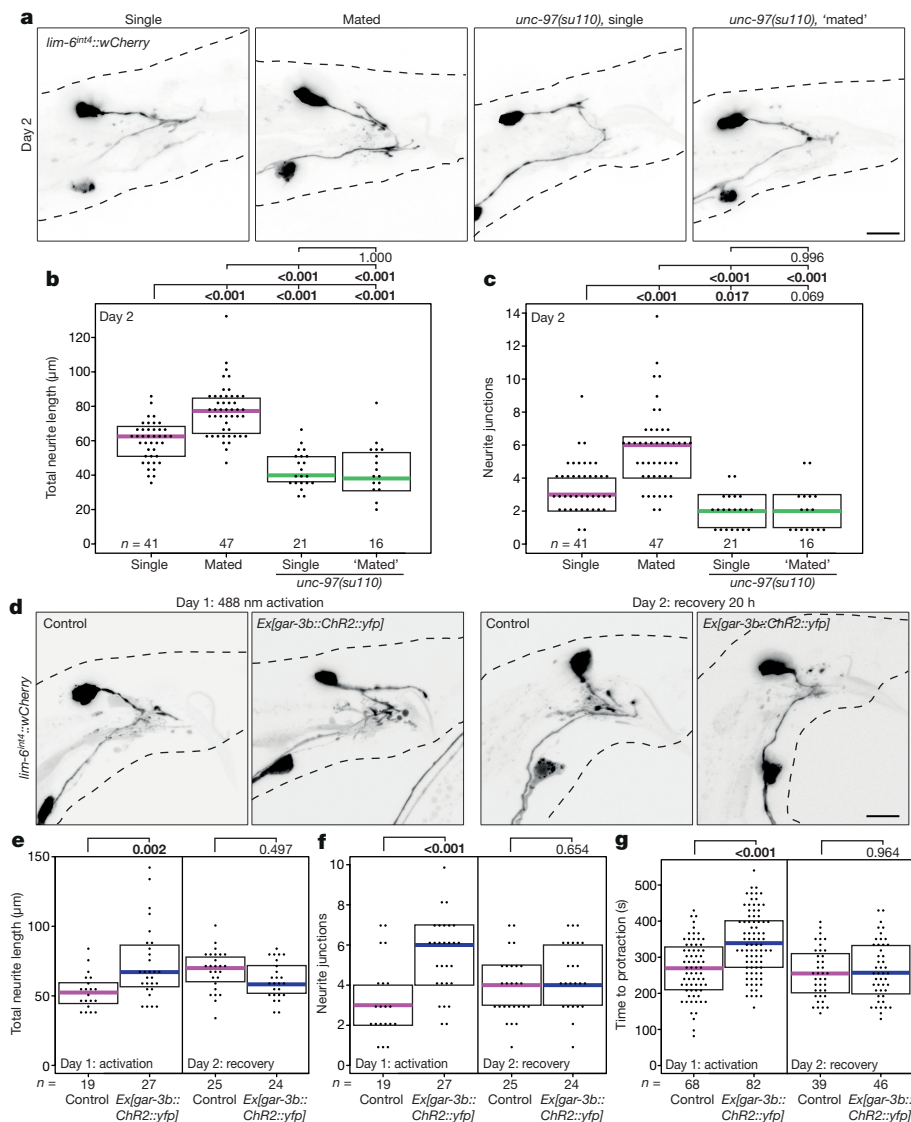
adulthood can rewire specific synaptic targets, supporting the notion that this remodelling can markedly alter connectivity within circuits and alter downstream behaviour.

Male spicule protraction into the hermaphrodite vulva is the most complex step of the male mating behaviour, involving coordination of cholinergic and GABAergic signalling^{16–18}. The balance of excitatory and inhibitory signalling is crucial for successful spicule insertion, which must be further coordinated with changes in sex muscle excitability in early adulthood^{13,19–21}. Day 1 and day 3 males are proficient at most steps of mating²⁰; however, in five-minute timed mating assays, day 3 males were significantly more likely than day 1 males to successfully complete mating with sperm transfer (*P* = 0.003; Extended Data Fig. 5a). We scored the spicule-related steps of mating (spicule prodding and spicule protraction) and found that day 1 males showed more spicule prodding attempts overall and a lower ratio of protraction to prodding attempts compared with day 3 males (Extended Data Fig. 5b, c), indicating that day 1 males are less capable than day 3 males of transitioning from spicule prodding to spicule protraction. This suggests that the morphological and functional plasticity of DVB in males may fine-tune and coordinate the defecation and spicule protraction circuits to increase mating success.

DVB neurites are experience- and activity-dependent

To determine whether DVB plasticity occurs in response to experience, we tested whether the act of mating itself altered DVB neuron morphology by exposing males to hermaphrodites for the first 48 h of adulthood. Single males housed with hermaphrodites showed significant increases in DVB neurite length and junctions compared to males housed alone (*P* < 0.001; Fig. 3a–c). *C. elegans* males housed with other males or in isolation can engage in mating-like behaviours, which may include spicule protraction. To minimize mating sensory input and self-mating behaviour, we analysed DVB neurite outgrowth in *pkd-2* (cation channel) mutant males²² and in genetically paralysed mutant males (*unc-97*)²³. *pkd-2* mutant males have reduced DVB neurite outgrowth at day 3, whereas *unc-97* mutant males have almost no DVB neurites at day 3 (Extended Data Fig. 4e–g); however, they can protract spicules in response to aldicarb (data not shown) and their neurites can be ectopically induced (Extended Data Fig. 5d–f). Paralysed males also show no change in neurite outgrowth when housed with hermaphrodites for 48 h (Fig. 3a–c). These results demonstrate that DVB neurite outgrowth is experience-dependent and is potentially driven by spicule protraction and activity of the postsynaptic spicule protraction circuit.

We next investigated whether activity of the postsynaptic targets of DVB contributes to DVB neurite outgrowth. Channelrhodopsin-



mediated activation of postsynaptic DVB targets (spicule neurons and muscle) resulted in immediate protraction of spicules¹⁶ (Fig. 2b, Supplementary Video 2). Repeated activation of the spicule protraction circuit caused a significant increase in DVB neurite length and junctions ($P = 0.002$ and $P < 0.001$, respectively; Fig. 3d–f, day 1), independent of GABA signalling (Extended Data Fig. 5d–f). Males exposed to repeated activation, but subsequently allowed to recover, had DVB neurites that were indistinguishable from those of controls, suggesting that neurite growth is dynamic and potentially reversible (Fig. 3d–f). Repeated activation of either spicule neurons or muscles separately demonstrated that activity in either can induce DVB neurite growth (Extended Data Fig. 5g–i).

We next tested whether activity-induced DVB neurites influence DVB neuron function and worm behaviour. We activated and recovered males in the same manner as above, and then used the aldicarb assay to analyse spicule protraction behaviour. Males at day 1 following repeated activation of the spicule protraction circuit showed a significant delay in the time to aldicarb-induced protraction ($P < 0.001$; Fig. 3g, day 1), implying that activity-induced neurites have a direct and immediate effect on DVB spicule function. Males that were exposed to repeated activation of the spicule protraction circuit but allowed to recover had spicule protraction indistinguishable from that of day 2 controls (Fig. 3g, day 2), indicating that induced behavioural changes are dynamic and repeated activation does not result in lasting protraction defects.

To test whether a reduction in circuit activity affects DVB neurites, we exposed males to exogenous GABA, expecting to silence the targets of GABAergic DVB signalling. This resulted in a reduction in DVB neurites (Extended Data Fig. 6a–c). To implicate spicule circuit inhibition more specifically, we silenced spicule protraction neurons and muscles with a histamine-gated chloride channel in day 5 males; this also reduced DVB neurites (Extended Data Fig. 6d–f). In summary, DVB neurites extend in response to the activity levels of the spicule protraction circuit, including postsynaptic targets of DVB.

Neurexin and neuroligin control DVB plasticity

DVB neurite outgrowth appears to be a form of morphological and functional plasticity that fine-tunes the excitatory and inhibitory balance for coordinated spicule protraction. Several synaptic molecules have been implicated in excitatory and inhibitory balance, including the synaptic adhesion molecule neurexin and its trans-synaptic binding partner neuroligin^{24–27}. Males with a deletion allele of the single *C. elegans* neuroligin orthologue *nlg-1* show increased DVB neurite outgrowth at day 3 compared to controls (Fig. 4a–c). The increase in DVB neurite outgrowth at day 3 was rescued by GFP-tagged NLG-1 expressed under its own promoter (Extended Data Fig. 7a–c), which was localized in a punctate pattern in numerous neurons and muscles of the male tail (Extended Data Fig. 8). *nlg-1* mutants displayed a spicule protraction phenotype that matches the expected phenotypes observed upon increased DVB branching (Fig. 4d). Expression of NLG-1 in the

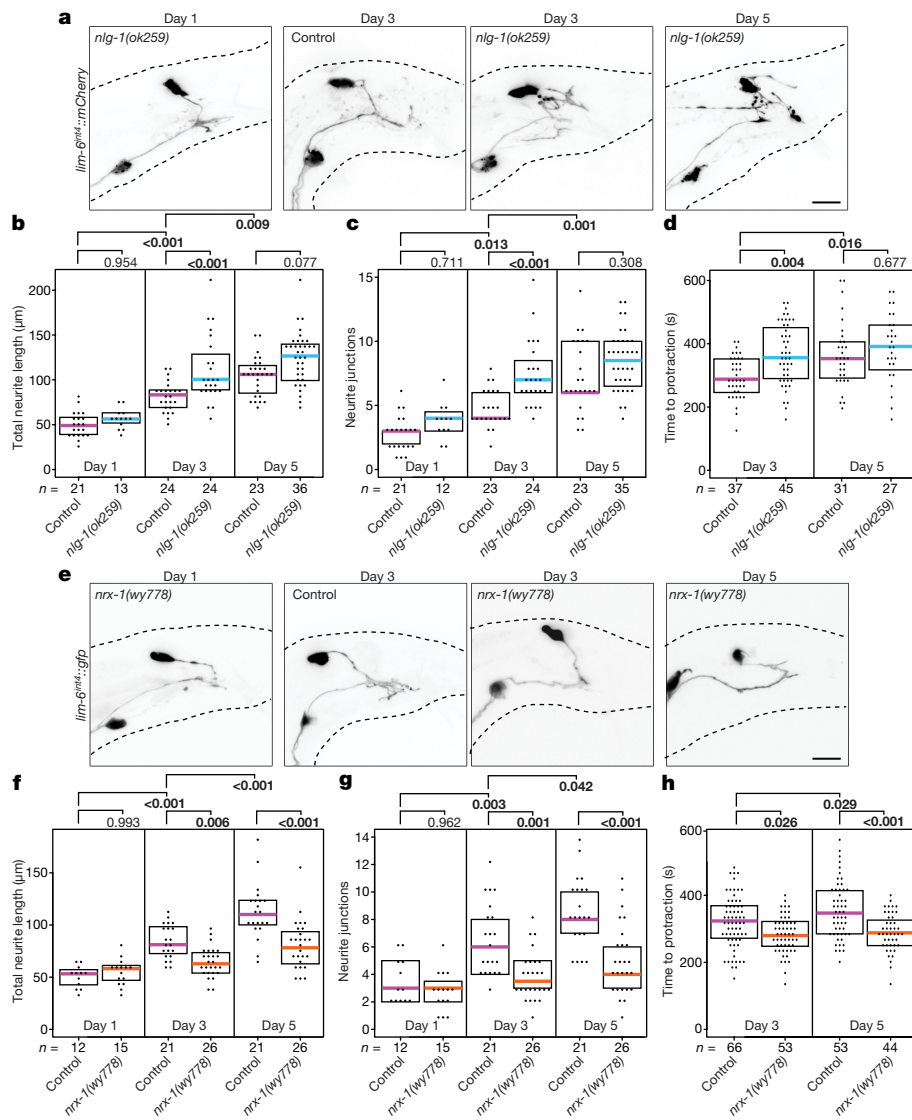


Figure 4 | Neuroigin and neuroligin influence DVB neurite outgrowth and spicule protraction behaviour. **a**, Confocal images of DVB (*lim-6^{int4}::wCherry*) in *nlg-1(ok259)* mutant and control male worms. **b**, **c**, Quantification of total neurite outgrowth (**b**) and number of neurite junctions (**c**) in *nlg-1(ok259)* and control male worms. **d**, Time to aldacar-induced spicule protraction in control and *nlg-1(ok259)* males. **e**, Confocal images of DVB (*lim-6^{int4}::gfp*) in *nrx-1(wy778)* mutant and control worms. **f**, **g**, Quantification of total neurite outgrowth (**f**) and number of neurite junctions (**g**) in *nrx-1(wy778)* and control worms. **h**, Time to aldacar-induced spicule protraction in control and *nrx-1(wy778)*. Scale bars, 10 µm. Dot represents one worm; magenta bar, median; boxes, quartiles. One-way ANOVA and post-hoc Tukey HSD. P values shown above plots, bold shows significance ($P < 0.05$).

DVB neuron, the SPC, PCA and PCB neurons, or the SPC neuron and spicule muscles did not rescue the *nlg-1* mutant phenotype, whereas expression in the spicule protractor and anal depressor muscles or in the spicule retractor muscles did rescue the phenotype (Extended Data Fig. 7d, e), indicating that NLG-1 contributes to DVB neurite outgrowth by functioning in multiple postsynaptic DVB muscles. Silencing the spicule protraction circuit in *nlg-1* mutant males at day 5 with *gar-3b::HisCl1* or overnight exposure to exogenous GABA resulted in no significant reduction in DVB neurite branching (Extended Data Fig. 7f, g). These results suggest that the *nlg-1* mutant phenotype cannot be explained by indirect alteration of the spicule circuit or more global perturbations in activity as a result of loss of NLG-1.

Unexpectedly, males with a deletion allele of *nrx-1* (which encodes the *C. elegans* orthologue of neuroligin)²⁸ displayed a significant reduction in neurite outgrowth at days 3 and 5, a phenotype opposite to the *nlg-1* mutant phenotype ($P = 0.006$ and $P < 0.001$, respectively; Fig. 4e–g). *nrx-1* mutants showed a corresponding decrease in time to aldacar-induced spicule protraction (Fig. 4h). The *nrx-1* locus produces both a long and short isoform²⁹, and two long isoform-specific mutant alleles recapitulated the null phenotype (Extended Data Fig. 9a–c). Repeated channelrhodopsin-mediated activation of the spicule protraction circuit failed to induce DVB neurites in *nrx-1* mutants (Extended Data Fig. 5d–f), indicating that the *nrx-1* phenotype is not explained solely by reduced circuit activity that could be envisioned to result from loss of NRX-1.

NRX-1 is broadly expressed throughout the *C. elegans* nervous system²⁹. Expression of the long isoform of NRX-1 in DVB using the *lim-6^{int4}* promoter resulted in rescue of the *nrx-1(wy778)* neurite outgrowth defect (Extended Data Fig. 9d, e). The long NRX-1 isoform still rescued the mutant phenotype even after deletion of the C-terminal PDZ binding motif, whereas the short NRX-1 isoform did not (Extended Data Fig. 9d, e). Overexpression of the long isoform of NRX-1 in wild-type male DVB neurons significantly increased DVB neurite length ($P = 0.047$) (Extended Data Fig. 9d, e), and when tagged with GFP, localized diffusely on the soma and neurites of DVB (Extended Data Fig. 9j). The reduction in time to aldacar-induced spicule protraction in *nrx-1* mutants was rescued by expression of the long isoform of NRX-1 in DVB, but overexpression of NRX-1 in wild-type worms did not change time to spicule protraction compared with control wild-type males (Extended Data Fig. 9f). These results indicate that the long isoform of NRX-1 is required in DVB for neurite outgrowth, which may extend the gene's role beyond its canonical function at synapses. Varying the levels of NRX-1 in DVB directly alters the extent of neurite outgrowth, and loss of NRX-1 in DVB reduces inhibition onto the spicule protraction circuit so that spicule protraction occurs more rapidly.

The exuberant DVB neurite branching phenotype of *nlg-1* mutants is completely suppressed by loss of NRX-1, and the increase in DVB neurite branching observed upon NRX-1 overexpression is not further enhanced by loss of NLG-1 (Extended Data Fig. 9g–i). Furthermore,

nrx-1(wy778);nlg-1(ok259) double null mutant males with NRX-1 expressed in DVB showed an increase in neurites, similar to *nlg-1* mutants (Extended Data Fig. 9g–i). Hence, restoration of NRX-1 expression in DVB with otherwise global loss of NRX-1 and NLG-1 recapitulates NLG-1 loss alone, suggesting that the *nlg-1* phenotype requires NRX-1 in DVB. GFP-tagged NRX-1 localized diffusely onto the membranes of soma and processes and did not appear to change between days 1 and 3 (Extended Data Fig. 9j). By contrast, expression of GFP-tagged NLG-1 decreased from days 1 to 3 in DVB-targeted muscles and neurons (Extended Data Fig. 8). Hence, NRX-1 appears to function cell-autonomously in DVB to promote DVB neurite outgrowth, whereas NLG-1 operates in postsynaptic partners of DVB to antagonize NRX-1-dependent growth. Decreases in NLG-1 expression may result in a reduction in the antagonistic relationship, thereby permitting more NRX-1-dependent neurite elaboration. Our demonstration of an antagonistic neurexin–neuroligin relationship that influences neurite outgrowth hints at a signalling process downstream of neurexin that is antagonized by neuroligin and is independent of neurexin's PDZ domain.

Finally, we tested whether manipulations that induce DVB neurites in males can also induce neurites in hermaphrodite DVB neurons. Activation of the anal depressor muscle (*gar-3b::ChR2::yfp*), loss of NLG-1, loss of NRX-1, or overexpression of NRX-1 in DVB had no effect on the axon morphology of hermaphrodite DVB neurons (Extended Data Fig. 10). Cell-autonomous sexual identity changes of either DVB or postsynaptic muscles using genetic manipulations of the sex-determination pathway also did not alter DVB morphology (see Methods). Thus, sexually dimorphic morphology and plasticity of the sex-shared DVB neuron seems to be non-autonomously instructed by male-specific circuit components.

Experience-dependent neuronal plasticity in the adult brain can include remodelling of dendrites and axons for behavioural adaptation or homeostatic maintenance of circuits. Our findings regarding male-specific DVB neurite outgrowth in *C. elegans* reveal the functional effect of morphological remodelling on circuits and behaviour. Through neurite outgrowth and rewiring of specific synapses, the DVB neuron undergoes a functional change that is likely to serve as an adaptive mechanism, perhaps translating experience into finer coordination of circuit activity and subsequent muscle contraction. These findings may have implications for the normal functions of neurexin and neuroligin in plasticity, and for the many human diseases associated with them.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 May; accepted 5 December 2017.

1. Fu, M. & Zuo, Y. Experience-dependent structural plasticity in the cortex. *Trends Neurosci.* **34**, 177–187 (2011).
2. Holtmaat, A. & Svoboda, K. Experience-dependent structural synaptic plasticity in the mammalian brain. *Nat. Rev. Neurosci.* **10**, 647–658 (2009).
3. Viturina, N. & Goda, Y. Cell biology in neuroscience: the interplay between Hebbian and homeostatic synaptic plasticity. *J. Cell Biol.* **203**, 175–186 (2013).
4. Yin, J. & Yuan, Q. Structural homeostasis in the nervous system: a balancing act for wiring plasticity and stability. *Front. Cell. Neurosci.* **8**, 439 (2015).
5. Lee, W. C. *et al.* Dynamic remodeling of dendritic arbors in GABAergic interneurons of adult visual cortex. *PLoS Biol.* **4**, e29 (2006).
6. Marić, S. A., Yamahachi, H., Meyer zum Alten Borgloh, S. & Gilbert, C. D. Large-scale axonal reorganization of inhibitory neurons following retinal lesions. *J. Neurosci.* **34**, 1625–1632 (2014).
7. Keck, T. *et al.* Loss of sensory input causes rapid structural changes of inhibitory neurons in adult mouse visual cortex. *Neuron* **71**, 869–882 (2011).

8. Jarrell, T. A. *et al.* The connectome of a decision-making neural network. *Science* **337**, 437–444 (2012).
9. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond. B* **314**, 1–340 (1986).
10. Reiner, D. J. & Thomas, J. H. Reversal of a muscle response to GABA during *C. elegans* male development. *J. Neurosci.* **15**, 6094–6102 (1995).
11. LeBoeuf, B. & Garcia, L. R. *Caenorhabditis elegans* male copulation circuitry incorporates sex-shared defecation components to promote intromission and sperm transfer. *G3* **7**, 647–662 (2017).
12. Pokala, N., Liu, Q., Gordus, A. & Bargmann, C. I. Inducible and titratable silencing of *Caenorhabditis elegans* neurons *in vivo* with histamine-gated chloride channels. *Proc. Natl Acad. Sci. USA* **111**, 2770–2775 (2014).
13. Garcia, L. R., Mehta, P. & Sternberg, P. W. Regulation of distinct muscle behaviors controls the *C. elegans* male's copulatory spicules during mating. *Cell* **107**, 777–788 (2001).
14. Hobert, O., Tessmar, K. & Ruvkun, G. The *Caenorhabditis elegans* *lim-6* LIM homeobox gene regulates neurite outgrowth and function of particular GABAergic neurons. *Development* **126**, 1547–1562 (1999).
15. Feinberg, E. H. *et al.* GFP reconstitution across synaptic partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* **57**, 353–363 (2008).
16. Liu, Y. *et al.* A cholinergic-regulated circuit coordinates the maintenance and bi-stable states of a sensory-motor behavior during *Caenorhabditis elegans* male copulation. *PLoS Genet.* **7**, e1001326 (2011).
17. Liu, Y., LeBoeuf, B. & Garcia, L. R. G_{αq}-coupled muscarinic acetylcholine receptors enhance nicotinic acetylcholine receptor signaling in *Caenorhabditis elegans* mating behavior. *J. Neurosci.* **27**, 1411–1421 (2007).
18. Jobson, M. A. *et al.* Spillover transmission is mediated by the excitatory GABA receptor LGC-35 in *C. elegans*. *J. Neurosci.* **35**, 2803–2816 (2015).
19. Garcia, L. R. & Sternberg, P. W. *Caenorhabditis elegans* UNC-103 ERG-like potassium channel regulates contractile behaviors of sex muscles in males before and during mating. *J. Neurosci.* **23**, 2696–2705 (2003).
20. Guo, X., Navetta, A., Gualberto, D. G. & Garcia, L. R. Behavioral decay in aging male *C. elegans* correlates with increased cell excitability. *Neurobiol. Aging* **33**, 1483.e5–1483.e23 (2012).
21. LeBoeuf, B. & Garcia, L. R. Cell excitability necessary for male mating behavior in *Caenorhabditis elegans* is coordinated by interactions between big current and ether-a-go-go family K⁺ channels. *Genetics* **190**, 1025–1041 (2012).
22. Barr, M. M. *et al.* The *Caenorhabditis elegans* autosomal dominant polycystic kidney disease gene homologs *lov-1* and *pkd-2* act in the same pathway. *Curr. Biol.* **11**, 1341–1346 (2001).
23. Hobert, O., Moerman, D. G., Clark, K. A., Beckerle, M. C. & Ruvkun, G. A conserved LIM protein that affects muscular adherens junction integrity and mechanosensory function in *Caenorhabditis elegans*. *J. Cell Biol.* **144**, 45–57 (1999).
24. Bang, M. L. & Owczarek, S. A matter of balance: role of neurexin and neuroligin at the synapse. *Neurochem. Res.* **38**, 1174–1189 (2013).
25. Maćkowiak, M., Mordalska, P. & Wędzony, K. Neuroligins, synapse balance and neuropsychiatric disorders. *Pharmacol. Rep.* **66**, 830–835 (2014).
26. Pizzarelli, R. & Cherubini, E. Alterations of GABAergic signaling in autism spectrum disorders. *Neural Plast.* **2011**, 297153 (2011).
27. Chih, B., Engelman, H. & Scheiffele, P. Control of excitatory and inhibitory synapse formation by neuroligins. *Science* **307**, 1324–1328 (2005).
28. Maro, G. S. *et al.* MADD-4/punctin and neurexin organize *C. elegans* GABAergic postsynapses through neuroligin. *Neuron* **86**, 1420–1432 (2015).
29. Haklai-Topper, L. *et al.* The neurexin superfamily of *Caenorhabditis elegans*. *Gene Expr. Patterns* **11**, 144–150 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank Q. Chen for generating transgenic strains and M. Gendrel for DVB promoter and reporter lines; P. Sengupta, T. G. Drivas, M. Oren-Suissa, and members of the Hobert laboratory for comments on the manuscript; L. R. Garcia and K. Shen for worm strains; and M. VanHoven and D. Colon-Ramos for plasmids. This work was supported by NIH grants from NINDS (M.P.H.:F32NS086285; O.H.:2R37NS039996). O.H. is a Howard Hughes Medical Institute investigator. Some strains were provided by the CGC, funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

Author Contributions M.P.H. and O.H. designed the experiments and wrote the manuscript. M.P.H. performed the experiments.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to M.P.H. (hartmicp@gmail.com).

METHODS

C. elegans strains. Wild-type strains were *C. elegans* variety Bristol, strain N2. Worms were grown at 23 °C on nematode growth medium (NGM) plates seeded with bacteria (*Escherichia coli* OP50) as a food source. All males contained either *him-8(e1489) IV* or *him-5(e1490) V* as indicated by strain. Male worms were picked at the fourth larval stage onto plates with ten other males (unless otherwise indicated), and allowed to moult into adults and age to the day indicated for each analysis or experiment.

Mutant alleles used in this study include: *him-8(e1489) IV*, *him-5(e1490) V*, *unc-31(e928) IV*, *nlg-1(ok259) X*, *nrx-1(ok1649) V*, *unc-119(ed3) III*, *nrx-1(wy778[unc-119(+)] V*, *lim-6(nr2073) X*, *pkd-2(pt8) IV*, *unc-97(su110) X*, *unc-25(e156) III*, *unc-49(e407) III*, *nrx-1(ok1649) V*, and *nrx-1(gk246237)*.

All transgenic strains used in this study are listed in Supplementary Table 1 ordered by Figures and Extended Data Figures. All plasmids were injected at 25 ng μl^{-1} with coinjection marker *ttx-3::gfp* or *ttx-3::wCherry* also at 25 ng μl^{-1} to generate extrachromosomal arrays (unless otherwise noted).

Cloning and constructs. To generate *lim-6^{int4}::wCherry* (pMG198) and *lim-6^{int4}::gfp* (pMG141), a 291-bp fragment of the *lim-6* fourth intron was amplified with primers adding BamHI to forward (CCCCGGATCCTTAGCCAGTTGCATAAATAT) and MscI to reverse (GGGGTGGCCACTAAGCTTCTTGCTAAATTC). This fragment was digested and ligated into a pPD95.75 vector with either GFP or codon-optimized mCherry (wCherry). Plasmids were injected at 5 ng μl^{-1} into a *pha-1(e2123)* mutant strain with *pha-1(+)* coinjection marker. Extrachromosomal arrays were integrated to yield *otIs541* and *otIs525*. *lim-6^{int4}* was found to express brightly in DVB, dimly in AVL and RIS, and dimly in about 70% of worms in PVT.

To generate *lim-6^{int4}::gfp::rab-3* (pMH1), *lim-6^{int4}* was PCR-amplified from pMG193 using primers forward GATGATACGTAACAACCTTGGAAATGA AATGGATCCTTAGCCAGTTGCATAAATATTAAGTCAAATG and reverse GAAACATACCTTTGGGTCCTTTGGCCACTAAGCTTCTTGCTAAATTCCTTTTGATTG, and cloned into DACR10 (a gift from D. Colon-Ramos) to replace the *ttx-3* promoter using restriction free cloning. The resulting plasmid was injected at 45 ng μl^{-1} with coinjection marker *ttx-3::gfp* also at 45 ng μl^{-1} . An extrachromosomal array was integrated to yield *otIs659*.

To generate *lim-6^{int4}::Chr2::yfp* (pMH17), *lim-6^{int4}* was PCR-amplified from pMH1 using primers forward CTAGATCAAACAAGTTTGACAAAAAAGCTT GCATGCCTGGATCCTTAG and reverse CACTTTGTACAAGAAAGCTGGGTC CTAAGCTTCTTGCTAAATTCCTTTTG, and cloned into pLR183 (*gar-3b::Chr2::yfp*, a gift from L. R. Garcia^{16,30}) to replace the *gar-3b* promoter using restriction free cloning.

To generate *lim-6^{int4}::BirA::nrx-1^{LONG}* (pMH27), *lim-6^{int4}* was PCR-amplified from pMH1 using primers forward GAAATGAAATAAAGCTTGCATGAG CTTGCATGCCTGGATCCTTAG and reverse CTTTGGGTCCTTTGGCCAAAT CCGGCTAAGCTTCTTGCTAAATTC, and cloned into pMO23³¹ (*srg-13::BirA::nrx-1*) to replace the *srg-13* promoter using restriction free cloning.

To generate *lim-6^{int4}::BirA::nrx-1^{SHORT}* (pMH41), the first exon of the *nrx-1* short isoform was PCR-amplified from N2 genomic DNA using primers forward GAAGT GGAGGTGGAGGCTCAGGTGTATTCTTGAGCATTTGCGTGGTG and reverse GTTGGAAAGCATGCGCAGAGAAGATCCAGTAGTCTCTCC GGACACATCATTC, and cloned into pMH27 to replace the first 23 exons of the long isoform of *nrx-1* using restriction free cloning.

To generate *lim-6^{int4}::BirA::nrx-1^{noPDZ}* (pMH44), the first exon of the *nrx-1* short isoform was PCR-amplified from N2 genomic DNA using primers forward CAACGGCCACAATGATGAGAAACGGAATGGGGTGGCATCT CGAGGAGCTCCCGAGATCTTCAGCGCTC and reverse CTACGAATGCTG AGCGCTGAAGATCTCGGAGCTCCTCGAGATTATGCCACCCATTCCC GTTTC, and cloned into pMH27 to delete the last 30 bp of *nrx-1* cDNA before the stop codon using restriction free cloning.

To generate *lim-6^{int4}::GFP::nrx-1* (pMH37), *eGFP* cDNA was PCR-amplified from pMH1 using primers forward CTATCGGAGCAGCATTAATACTAGGCA TTTGGCTCAAAAAGACTGTTACG and reverse CGACGATGAC GTAACAGTCTTTTGTAGCCAAATGCCTAGTATTGAATG, and cloned into pMH27 to replace *birA* cDNA using restriction free cloning.

To generate *lim-6^{int4}::nlg-1::gfp-1-10* (pMH18), *lim-6^{int4}* was PCR-amplified from pMH1 using primers forward CAAGCTTGCATGCGCGGCCGACAGCTT GCATGCCTGGATCCTTAG and reverse GTCCTTTGGCCAATCCCGGGGATCT AAGCTTCTTGCTAAATTCCTTTTG, and cloned into MVC6 (*gpa-6::nlg-1::gfp-1-10*, a gift from M. VanHoven) to replace the *gpa-6* promoter using restriction free cloning.

To generate *gar-3b::nlg-1::gfp11* (pMH20), the *gar-3b* promoter was PCR-amplified from pLR183 using primers forward CAAGCTTGCATGCGCGGCCG CACCATAAGCATCATGAGCAACATCTCCACTTCTCGTGAGC and reverse GTCCTTTGGCCAATCCCGGGGATGATTAATAATGTGCAGGAGGAGTA

ATAATGGTGTATGT, and cloned into MVC12 (*flp-18p::nlg-1::gfp11*, a gift from M. VanHoven) to replace the *flp-18* promoter using restriction free cloning.

To generate *lim-6^{int4}::nlg-1::gfp11* (pMH8), *lim-6^{int4}* was PCR-amplified from pMH1 using primers forward CAAGCTTGCATGCGCGGCCGACAGCT TGCATGCCTGGATCCTTAG and reverse GTCCTTTGGCCAATCCCGGGGAT CTAAGCTTCTTGCTAAATTCCTTTTG, and cloned into MVC6 to replace the *gar-3b* promoter using restriction free cloning.

To generate *flp-13::nlg-1::gfp11* (pMH23), the *flp-13* promoter was PCR-amplified from N2 genomic DNA using primers forward CAAGCTTGCATGCGCGGC CGCAGCAGTGACGTCATCTTGTTCG and reverse GTCCTTTGGCCAATC CCGGGGATAAATTTGTGCCTCCTGATGCTG, and cloned into pMH20 to replace the *gar-3b* promoter using restriction free cloning.

To generate *unc-103E::nlg-1::gfp11* (pMH21), the *unc-103E* promoter was PCR-amplified from N2 genomic DNA using primers forward CAAGCTTGCATGC GCGGCCGACTCGCGGTGCCCAAAGGTAGGTATTGACGTATCTCC and reverse GTCCTTTGGCCAATCCCGGGGATTACCACCACCACCACAAC CACCGATCGACGAC, and cloned into pMH20 to replace the *gar-3b* promoter using restriction free cloning.

To generate *unc-103F::nlg-1::gfp11* (pMH25), the *unc-103F* promoter was PCR-amplified from N2 genomic DNA using primers forward CAAGCTTGCA TGCGCGGCCGACCACGCTGCTAAGGGATGCCTTAGCTC and reverse GTCCTTTGGCCAATCCCGGGGATGACATTGCCACGTGGTTGTGTGTG, and cloned into pMH20 to replace the *gar-3b* promoter using restriction free cloning.

To generate *lim-6^{int4}::HisC11::gfp* (pMH3), the *lim-6^{int4}* promoter was PCR-amplified from N2 genomic DNA using primers forward GCATGCGCGGCCGCA CTGACTGGGCCGCGCGGATCCTTAGCCAGTTG and reverse CAATCCCGGG GATCCTCTAGAGCGCGCCCTAAGCTTCTTGCTAAATTC, and cloned into pNP471 to replace the *rig-3* promoter using restriction free cloning.

To generate *gar-3b::HisC11::gfp* (pMH28), the *gar-3b* promoter was PCR-amplified from pMH20 genomic DNA using primers forward CTTGCAT GCGCGGCCGACTGACTGGGCCGCGCCATAAGCATCATGAGCAACATC TC and reverse CAATCCCGGGGATCCTCTAGAGCGCGCCAAAGCTGG GTCGATTAATAAATGTGAC, and cloned into pMH3 to replace the *lim-6^{int4}* promoter using restriction free cloning.

Microscopy. Worms were anaesthetized using 100 mM sodium azide (NaN₃) and mounted on a pad of 5% agar on glass slides. Worms were analysed by Nomarski optics and fluorescence microscopy, using a Zeiss 880 confocal laser-scanning microscope. Multidimensional data were reconstructed as maximum intensity projections using Zeiss Zen software. Puncta were quantified by scanning the original full Z-stack for distinct dots in the area overlapping with the processes of the DVB neuron. Figures were prepared using Adobe Photoshop CS6 and Adobe Illustrator CS6.

Neurite tracing. Confocal Z-stacks were opened using FIJI, and loaded into the Simple Neurite Tracer plugin³². The primary neurite of DVB was traced from the centre of the cell soma to the point where the axon projects ventrally and then turns anteriorly, at the final branch point before it becomes a single process. Neurites were added by tracing off of this primary neurite, including all neurites emanating posterior of the last branch point. The simple neurite tracer plugin was used to analyse the skeletons for neurite length, which were summed to calculate total neurite length, and the number of neurite junctions (a proxy for the number of neurite branches).

Cell ablation. We performed laser ablations using a MicroPoint Laser System Basic Unit (N2 pulsed laser (dye pump), ANDOR Technology) attached to a Zeiss Axioplan 2IE widefield microscope (objective EC Plan-Neofluar 100 \times /1.30 Oil M27). This laser delivers 120 μJ of 337-nm energy with a 3-ns pulse length. Ablations were performed as previously described³³, with pulse repetition rates of \sim 15 Hz. Cell identification was performed with GFP or Cherry markers. Ablations were performed at the days of adulthood indicated, and worms were analysed \sim 20 h later. Mock-ablated worms were placed on same slide under the microscope but were not ablated, and were allowed to recover in a similar manner. Before relevant assays were performed (spicule protraction or aldicarb assays), worms were analysed for loss of cell fluorescence under a dissecting scope. When possible, after assays, worms were mounted on glass slides and analysed under a microscope to validate that cell ablation was successful.

Aldicarb spicule protraction assay. Aldicarb was added to warm liquid NGM agar medium to a final concentration of 5 mM and poured into plates. Worms were picked 12 or fewer at a time onto aldicarb plates and observed for spicule protractions longer than 5 s, when the time was recorded for each worm¹³.

Mating assay. L4 male worms were picked and singled onto plates. Non-mated males were left individually on plates, whereas mated males had 10 *unc-31(e928)* hermaphrodites added to their plates. We exposed males to uncoordinated hermaphrodites (*unc-31/CAPS*), to ensure a successful mating experience. Following 48 h of being housed either individually or with 10 hermaphrodites, all

mated plates were checked for fluorescent progeny to ensure successful mating had occurred, and then mated and non-mated (individually housed) males were subjected to confocal microscopy. *C. elegans* males housed with other males or in isolation can engage in mating-like behaviours, which may include spicule protraction. To minimize mating sensory input and self-mating behaviour, we also analysed DVB neurite outgrowth in males with mutation in *pkd-2*²² and in males genetically paralysed by a mutation in *unc-97/LIMS1* (affects body wall muscle ultrastructure)²³.

Mating behaviour assay. Mating assays were based on procedures described previously^{34,35}. Males were picked at the L4 stage and kept apart from hermaphrodites. One male was transferred to a plate covered with a fresh OP50 lawn containing 15 adult *unc-31(e28)* hermaphrodites. Day 1 males were counted as less than 18 h after L4 moult. Males were observed for 5 min from the time of first contact with a hermaphrodite or until they ejaculated, whichever came first. Males were scored for their ability to prod the vulva, protract spicules, and transfer sperm. Mating success was calculated as $100 \times$ the number of males that transferred sperm successfully divided by the total number of males tested. The number of attempts at prodding was calculated by summing attempts at prodding for each male. The protraction:prodding ratio was calculated by dividing the number of spicule protractions by the number of attempts at prodding for each male.

Synapse visualization. GRASP plasmid construction is described above. For visualization of synaptic connections between DVB and neurons and muscles downstream of DVB that form the spicule protraction circuit, we injected *lim-6^{int4}::nlg-1::gfp1-10* (pMH18) to label the presynaptic DVB, together with *gar-3b::nlg-1::gfp11* (pMH20) to label the postsynaptic SPC and spicule protractor muscles. Plasmids were injected together at $25 \text{ ng } \mu\text{l}^{-1}$ with the coinjection marker *ttx-3::gfp* (also at $25 \text{ ng } \mu\text{l}^{-1}$) to generate extrachromosomal arrays. For visualization of synaptic connections between DVB and neurons and muscles downstream of DVB that express *flp-13*, we injected *lim-6^{int4}::nlg-1::gfp1-10* (pMH18) to label the presynaptic DVB together with *flp-13::nlg-1::gfp11* (pMH23) to label the postsynaptic spicule retractor muscles. Plasmids were injected together at $25 \text{ ng } \mu\text{l}^{-1}$ with the coinjection marker *ttx-3::gfp* (also at $25 \text{ ng } \mu\text{l}^{-1}$) to generate extrachromosomal arrays. Synapses between DVB and spicule retractor muscles were not reported in electron microscopy of an 'old male'³⁶, possibly owing to the observed decrease in these synapses after day 1; alternatively, these synapses may have been characterized as one of several 'unknown' connections of DVB⁸. The *flp-13* promoter also labels CP6 in males, which has few synapses with DVB that were located in the electron microscopy reconstruction anterior to the DVB neurites, and the branched parts of the axons of DVB and CP6 appear not to make contact (Extended Data Fig. 4h).

Spicule activation assay with channelrhodopsin. All-*trans* retinal was added to LB/OP50 medium and coated over the entire plate at a final concentration of 0.1 mM. We obtained strains expressing channelrhodopsin under the *gar-3b* promoter^{17,30,37} labelling spicule protraction neurons and muscles, under the *unc-103E* promoter labelling spicule retractors and anal depressor muscles, and under the *unc-103F* promoter labelling spicule neurons SPC, PCA, and PCB¹⁶ (gifts from L. R. Garcia). Worms were incubated overnight on retinal plates before all assays involving channelrhodopsin-containing strains. For the spicule protraction assay, male worms on retinal plates were individually subjected to 488-nm light for 10 s, three times with 30 s between trials, on a Nikon eclipse E400 microscope. Obvious spicule muscle contraction for any of the three trials was recorded as a response. Videos were recorded using a mounted Exo Labs Focus camera. For the activation protocol, male worms on retinal plates were subjected to alternating 488-nm light three times (15 s light/15 s dark) on a Leica M165 FC dissecting scope, repeated every 45 min for 4.5 h. Worms were then subjected to confocal microscopy or aldicarb behavioural assay. Controls for neurite outgrowth and aldicarb behaviour were performed on males under the same conditions but not exposed to the channelrhodopsin cofactor all-*trans* retinal (Extended Data Fig. 5j–l). For recovery, worms were placed in the dark for ~20 h after the activation protocol, then subjected to the same analysis. A small number of individual males subjected to confocal imaging before and after activation, or after activation and following recovery, demonstrated addition of neurites following activation, and removal of neurites following recovery; however, the difficulty of this analysis precluded quantification.

Neuronal silencing with histamine chloride channel (HisCl1). Control or transgenic worms were picked onto normal NGM plates seeded with OP50 at the L4 stage, then picked the evening before the indicated day of analysis onto 10 mM histamine or control plates with OP50 bacteria as a food source. For *gar-3b::HisCl1* silencing assays, males were left on histamine or control plates overnight then subjected to confocal microscopy the following morning. For *lim-6^{int4}::HisCl1* defecation analysis, males were picked onto histamine plates, allowed to adjust for 5 min and then analysed for defecation behaviour. Histamine plates were prepared as previously described¹².

Defecation assay. Males were placed on control or 10 mM histamine plates with food on the day of analysis, allowed to explore for 5 min, and then observed for 10–12 min on a low magnification Leica MZ8 light dissecting microscope.

Expulsion steps were recorded for the time between consecutive expulsions, and the presence of spicule protraction within 3 s before or after expulsion. The percentage of expulsion steps associated with spicule protraction was calculated for each male. The time between consecutive expulsion steps was calculated by averaging all times recorded between consecutive expulsions for each male.

Exogenous GABA exposure. Males were picked onto normal NGM plates seeded with OP50 at the L4 stage, then picked before the day of analysis onto 30 mM GABA¹⁰ or control plates seeded with OP50 and left overnight, and then subjected to confocal microscopy. For 3-day GABA exposure, males were picked onto 30 mM GABA or control plates seeded with OP50, left for 3 days and then subjected to confocal microscopy.

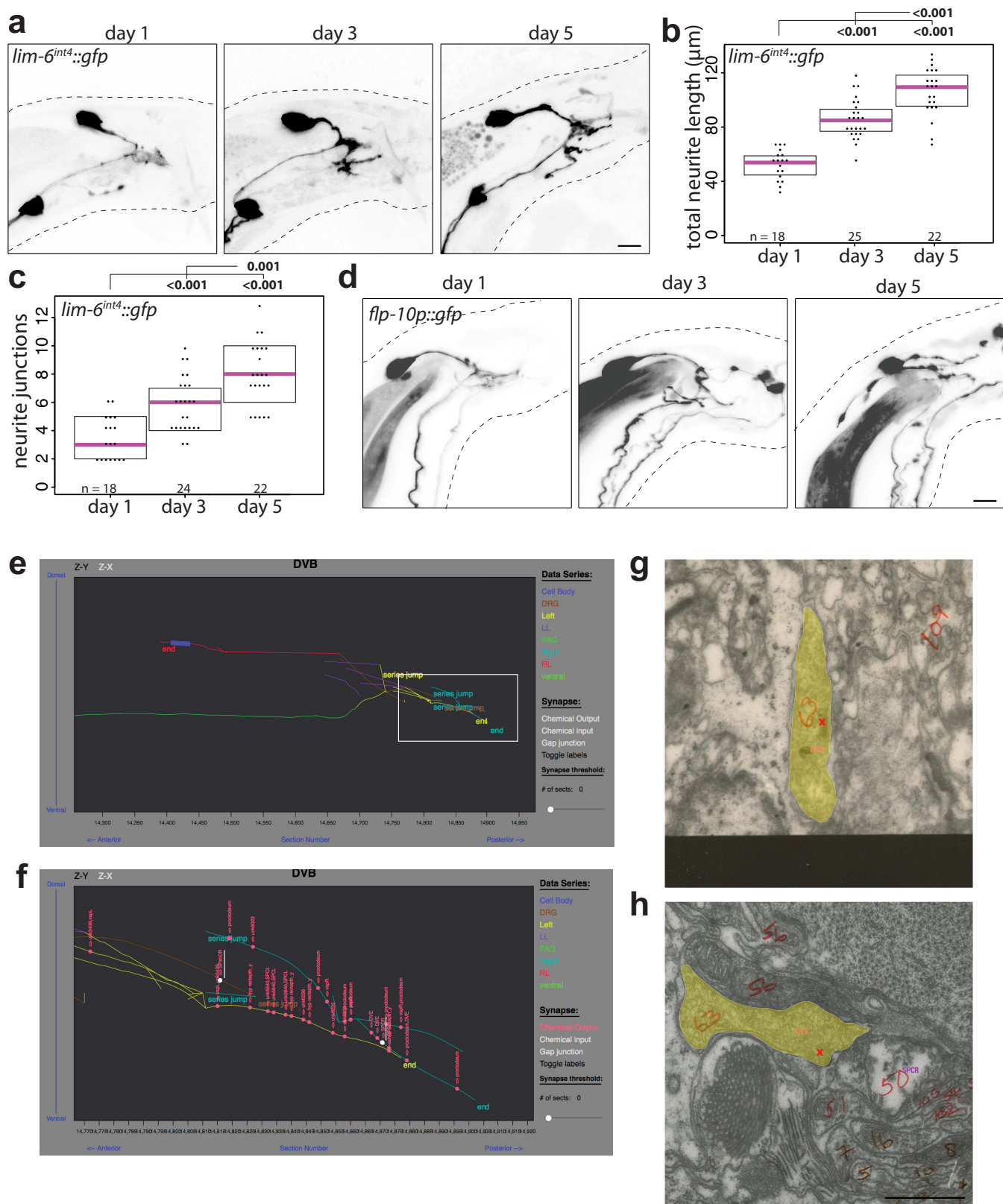
Measurement of fluorescence intensity. To quantify the fluorescence intensity of *nlg-1p::nlg-1::gfp*, a stack of images was acquired using confocal microscopy with the same acquisition parameters between samples (objective, pixel size, laser intensity, pinhole size, and PMT settings). The fluorescence intensity mean was obtained using ZEN Black software. For the dorsal spicule muscles, the muscles were outlined and the cross-section with the highest mean was recorded. Dorsal spicule muscles include the gubernacular retractor, gubernacular erector, anterior oblique, and anal depressor, which could be outlined easily, whereas the spicule protractor could not always be observed in males after day 1. For the pre-anal ganglion and DVB or background, a pre-defined circle was used to outline the region of interest, and the cross section with the highest mean was recorded. The ratio of fluorescence intensity was calculated by dividing the mean of the dorsal spicule muscles (arbitrary units) by the mean of the DVB or background (arbitrary units) or by dividing the mean of the pre-anal ganglion by the mean of the DVB or background (arbitrary units).

Cell-autonomous changes in sexual identity. We tested cell-autonomous changes in the sexual identity of DVB (*lim-6^{int4}* promoter) and muscles (*myo-3* promoter) by expressing either the cDNA of *fem-3* in hermaphrodites to masculinize each tissue, or the cDNA of *tra-2^{intracellular domain}* in males to feminize each tissue^{38–40}. In males with feminized DVB or muscles, we observed no suppression of DVB neurites, and in hermaphrodites with masculinized DVB or muscle, we observed no induction of DVB neurites.

Statistics and reproducibility. We performed two-tailed Student's *t*-test or one-way ANOVA with post-hoc Tukey HSD test using *R* and *RStudio*; *P* values are shown on each graph. No statistical methods were used to predetermine sample size, and the experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. Number of independent biological replicates: Fig. 1b–d, 7; Fig. 2a–c, h–k, 3 or more; Fig. 2e, f, 2 or more; Figs 3a–g, 3 or more; Figs 4a–h, 3 or more; Extended Data Fig. 1a–c, 4 or more; Extended Data Fig. 1d, 2 or more; Extended Data Fig. 2a, b, 4 or more; Extended Data Fig. 2c–h, 2 or more; Extended Data Fig. 3a–c, 3 or more; Extended Data Fig. 3d–f, 2 or more; Extended Data Fig. 4a–c, h, 2 or more; Extended Data Fig. 4d–g, 3 or more; Extended Data Fig. 5a–f, 4 or more; Extended Data Fig. 5g–l, 2 or more; Extended Data Fig. 6a–f, 2 or more; Extended Data Fig. 7a–h, 3 or more; Extended Data Fig. 8a–c, 3 or more; Extended Data Fig. 8d–f, 2 or more; Extended Data Fig. 9b–i, 3 or more; Extended Data Fig. 9j, 2 or more; Extended Data Fig. 10a–c, 3 or more.

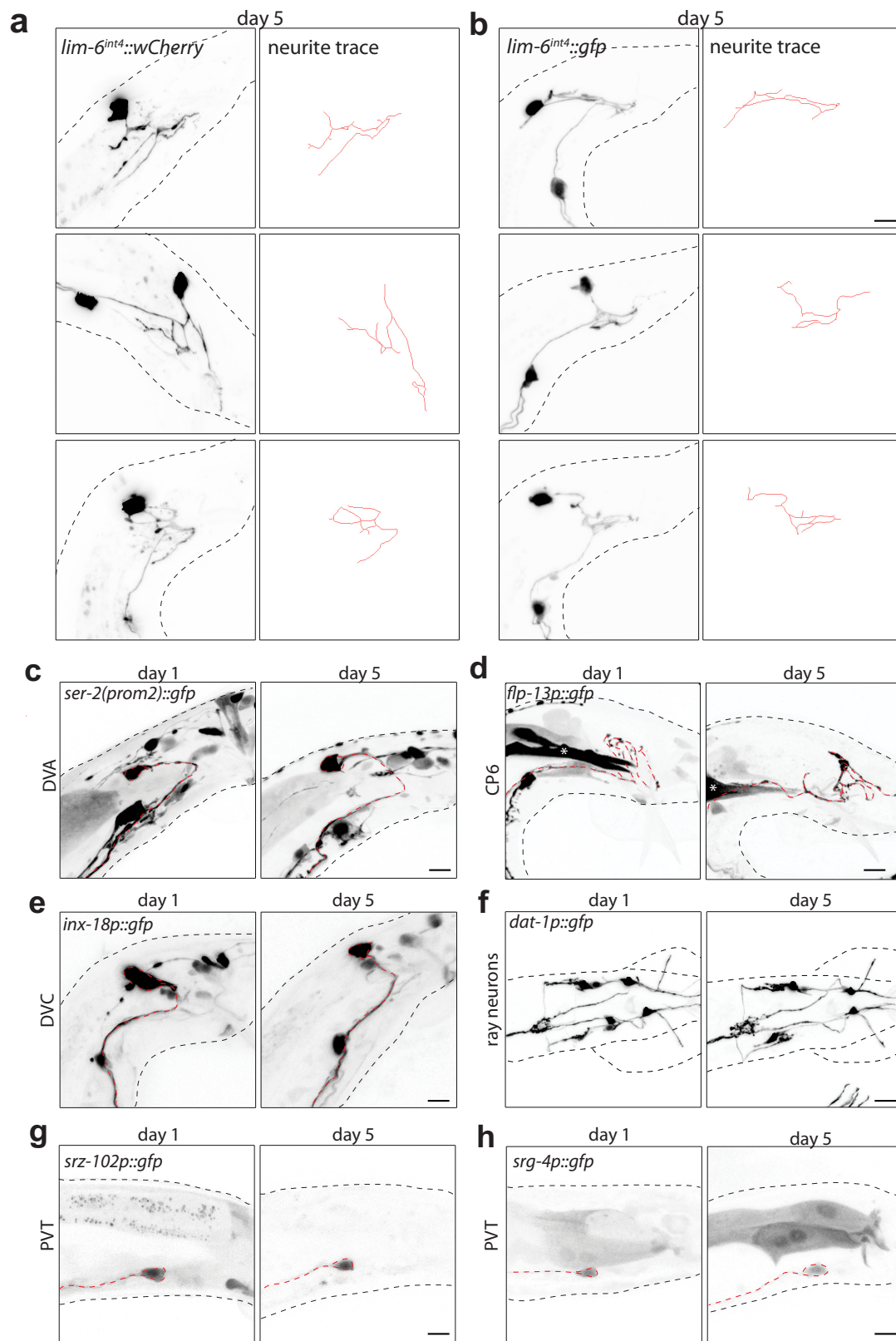
Data availability. The data that support the findings of this study are available from the corresponding author upon reasonable request.

- LeBoeuf, B., Correa, P., Jee, C. & Garcia, L. R. *Caenorhabditis elegans* male sensory-motor neurons and dopaminergic support cells couple ejaculation and post-ejaculatory behaviors. *eLife* **3**, e02938 (2014).
- Oren-Suissa, M., Bayer, E. A. & Hobert, O. Sex-specific pruning of neuronal synapses in *Caenorhabditis elegans*. *Nature* **533**, 206–211 (2016).
- Longair, M. H., Baker, D. A. & Armstrong, J. D. Simple Neurite Tracer: open source software for reconstruction, visualization and analysis of neuronal processes. *Bioinformatics* **27**, 2453–2454 (2011).
- Fang-Yen, C., Gabel, C. V., Samuel, A. D., Bargmann, C. I. & Avery, L. Laser microsurgery in *Caenorhabditis elegans*. *Methods Cell Biol.* **107**, 177–206 (2012).
- Garcia, L. R., LeBoeuf, B. & Koo, P. Diversity in mating behavior of hermaphroditic and male–female *Caenorhabditis* nematodes. *Genetics* **175**, 1761–1771 (2007).
- Hilliard, M. A., Bargmann, C. I. & Bazzicalupo, P. C. *elegans* responds to chemical repellents by integrating sensory inputs from the head and the tail. *Curr. Biol.* **12**, 730–734 (2002).
- Sulston, J. E., Albertson, D. G. & Thomson, J. N. The *Caenorhabditis elegans* male: postembryonic development of nongonadal structures. *Dev. Biol.* **78**, 542–576 (1980).
- Correa, P., LeBoeuf, B. & Garcia, L. R. C. *elegans* dopaminergic D2-like receptors delimit recurrent cholinergic-mediated motor programs during a goal-oriented behavior. *PLoS Genet.* **8**, e1003015 (2012).
- White, J. Q. et al. The sensory circuitry for sexual attraction in *C. elegans* males. *Curr. Biol.* **17**, 1847–1857 (2007).
- Lee, K. & Portman, D. S. Neural sex modifies the function of a *C. elegans* sensory circuit. *Curr. Biol.* **17**, 1858–1863 (2007).
- Mowrey, W. R., Bennett, J. R. & Portman, D. S. Distributed effects of biological sex define sex-typical motor behavior in *Caenorhabditis elegans*. *J. Neurosci.* **34**, 1579–1591 (2014).



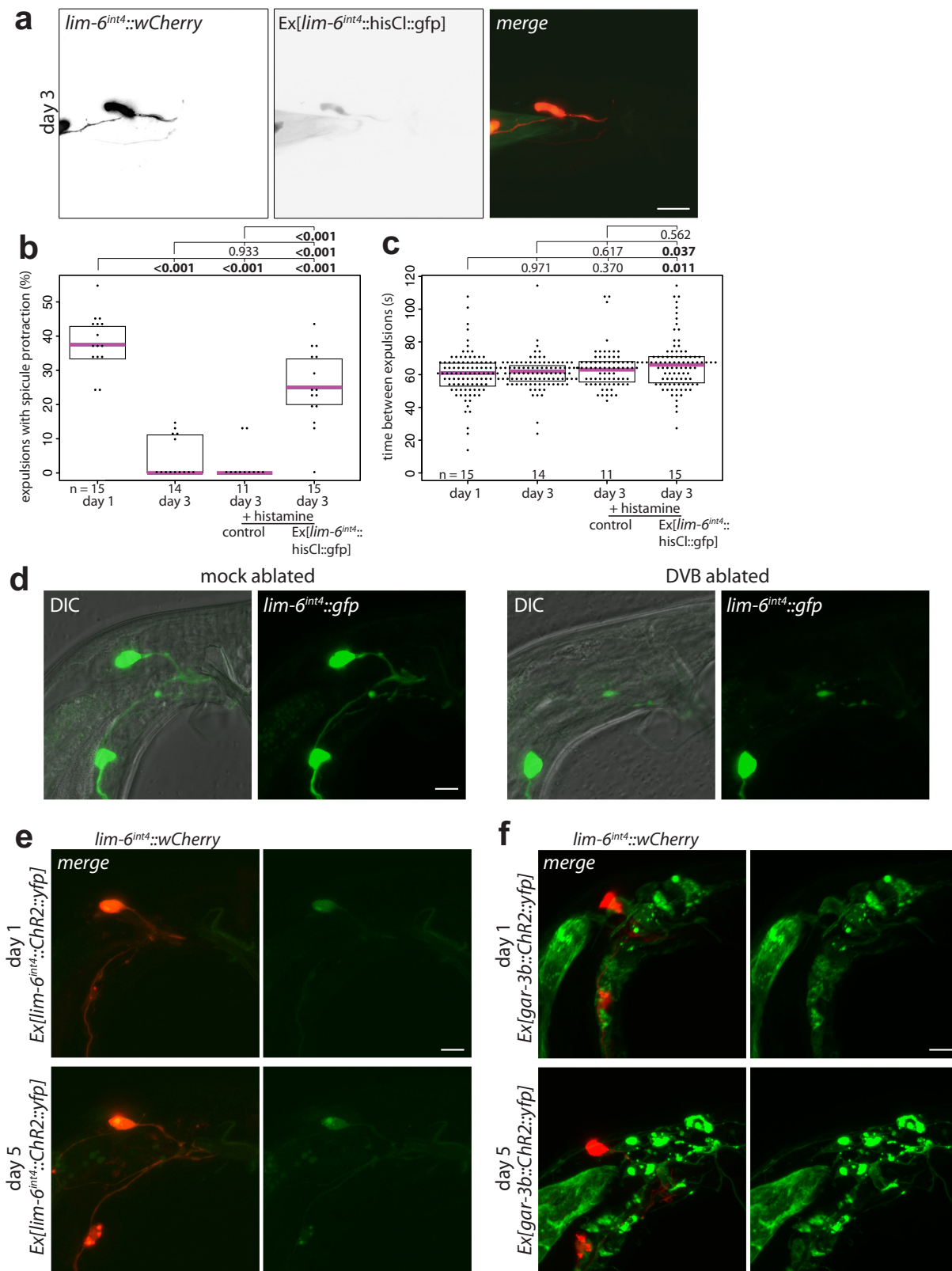
Extended Data Figure 1 | Progressive neurite outgrowth in DVB in adulthood. **a–c**, DVB neuron visualized with *lim-6^{int4}::gfp* at days 1, 3, and 5 in adult males (**a**) and quantification of total neurite length (**b**) and number of neurite junctions (**c**) (dot represents one worm; magenta bar, median; boxes, quartiles; one-way ANOVA and post-hoc Tukey HSD, *P* values shown above plots, bold shows significance (*P* < 0.05)). **d**, DVB neurite outgrowth visualized with *flp-10::gfp* in males at days 1, 3, and 5 of adulthood (*n* > 10, scale bars, 10 μm). **e**, Tracing reconstruction of male

DVB from electron micrograph sections compiled by <http://wormwiring.org> showing DVB neurites. **f**, Inset of DVB neurites showing presynaptic specializations identified in electron micrograph sections shown in pink. **g, h**, Electron micrograph section showing DVB pseudo-coloured yellow with presynaptic specialization indicated with red x with SPCR (Image Right1200, Section 14871) (**g**) and spicule sheath (Image N2YDRG1175, Section 14816) (**h**), shown in white in inset panel. Scale bars, 1 μm .



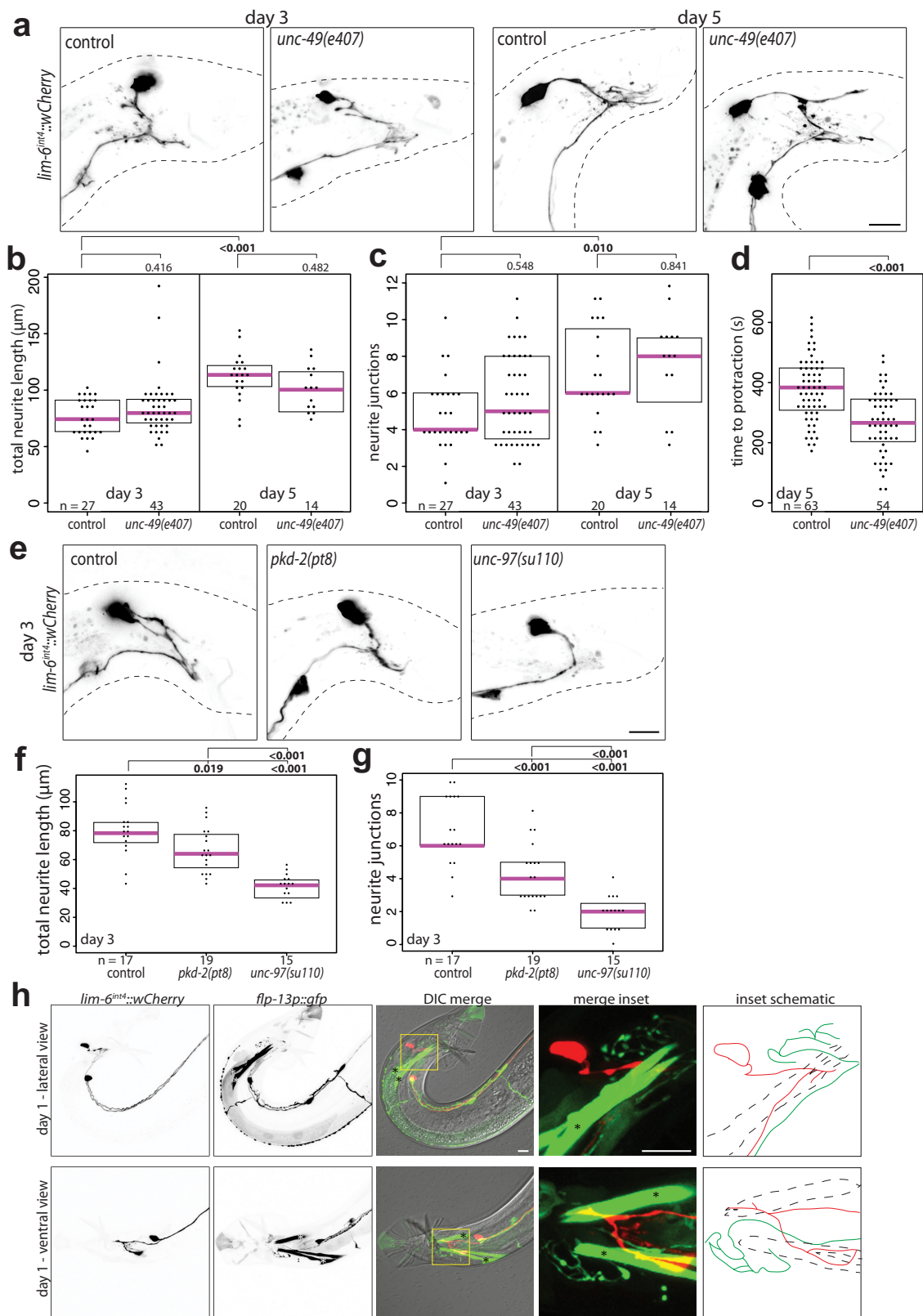
Extended Data Figure 2 | DVB neurite outgrowth in adult male *C. elegans* is stochastic and other neurons in the male tail do not show progressive neurite outgrowth in adulthood. a, b, DVB neurites at day 5 visualized with *lim-6^{int4}::wCherry* (a) or *lim-6^{int4}::gfp* (b) ($n > 10$ for each). DVB posterior neurites were traced through confocal stacks using Simple Neurite Tracer⁴ plugin. c, DVA neuron visualized with *ser-2(prom2)::gfp* ($n = 5$) (red dashed line indicates axon of relevant neuron). d, DVC neuron

visualized with *inx-18p::gfp* ($n = 5$). e, CP6 neuron visualized with *flp-13::gfp* (cell soma not shown) ($n = 5$). f, Ray neurons visualized with *dat-1::gfp* (ventral view) ($n = 5$). g, h, PVT neuron visualized with *srz-102p::gfp* ($n = 5$) (g) and *srg-4p::gfp* ($n = 5$) (h) at day 1 and day 5. Axons of indicated neurons highlighted by red dashed lines. Scale bars, 10 μ m.



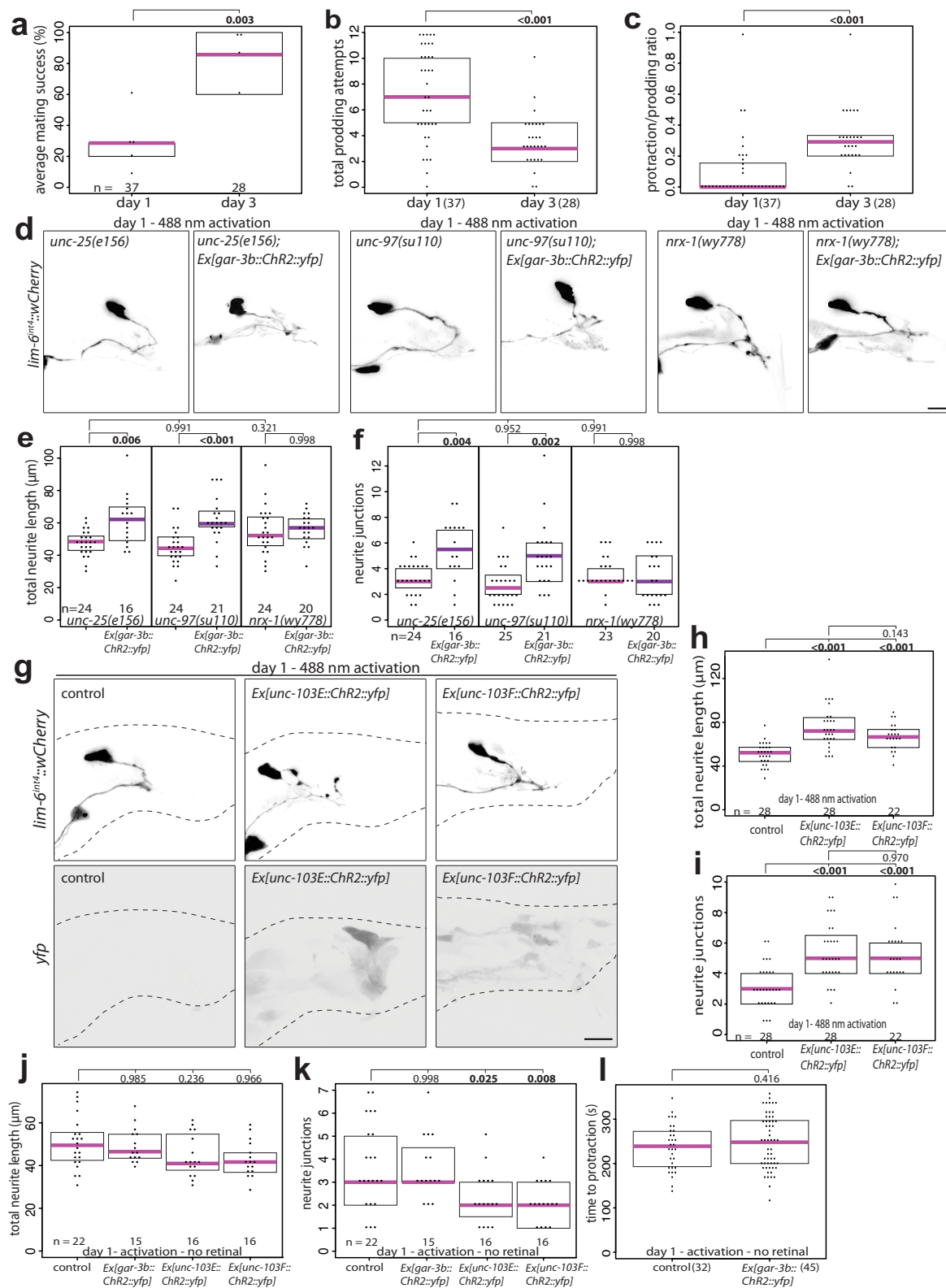
Extended Data Figure 3 | DVB inhibits expulsion-associated spicule protraction at day 3. Laser ablation of DVB and channelrhodopsin expression in DVB and spicule protraction circuit. **a**, Confocal images of male worm with *lim-6^{int4}::wCherry* and *lim-6^{int4}::HisCl::gfp* at day 3. **b**, Quantification of the percentage of expulsion steps with spicule protraction for day 1 control, day 3 control, day 3 control + histamine, and day 3 *lim-6^{int4}::HisCl::gfp* + histamine males. **c**, Time between consecutive expulsion steps for day 1 control, day 3 control, day 3 control + histamine, and day 3 *lim-6^{int4}::HisCl::gfp* + histamine males

(+ histamine is on 10 mM histamine plates; dot represents one worm; magenta bar, median; boxes, quartiles; one-way ANOVA and post-hoc Tukey HSD, *P* values shown above plots, bold shows significance (*P* < 0.05)). **d**, Confocal images of male worms with or without laser ablation of DVB at day 1 or 2, visualized with *lim-6^{int4}::gfp*. **e**, Confocal images of DVB (*lim-6^{int4}::wCherry*) expressing channelrhodopsin at day 1 and 5, *Ex[lim-6^{int4}::ChR2::yfp]*. **f**, Confocal images of DVB (*lim-6^{int4}::wCherry*) and spicule circuit expressing channelrhodopsin at day 1 and 5, *Ex[gar-3b::ChR2::yfp]*. *n* > 10 for **d–f**. Scale bars, 10 μ m.



Extended Data Figure 4 | DVB neurite outgrowth in *unc-49*, *pkd-2* and *unc-97* mutant males. *flp-13::gfp* labels CP6 and spicule retractor muscles. a–c, Confocal images (a) and quantification of total neurite outgrowth (b) and number of neurite junctions (c) in control and *unc-49(e407)* males at days 3 and 5. d, Time to spicule protraction on aldicarb at day 5 for control and *unc-49(e407)* males. e–g, Confocal images (e) and quantification of total neurite outgrowth (f) and number of neurite junctions (g) in control, *pkd-2(pt8)*, and *unc-97(su110)* males at day 3. h, Confocal images of male worms with *lim-6^{int4}::wCherry*, *flp-10p::gfp*,

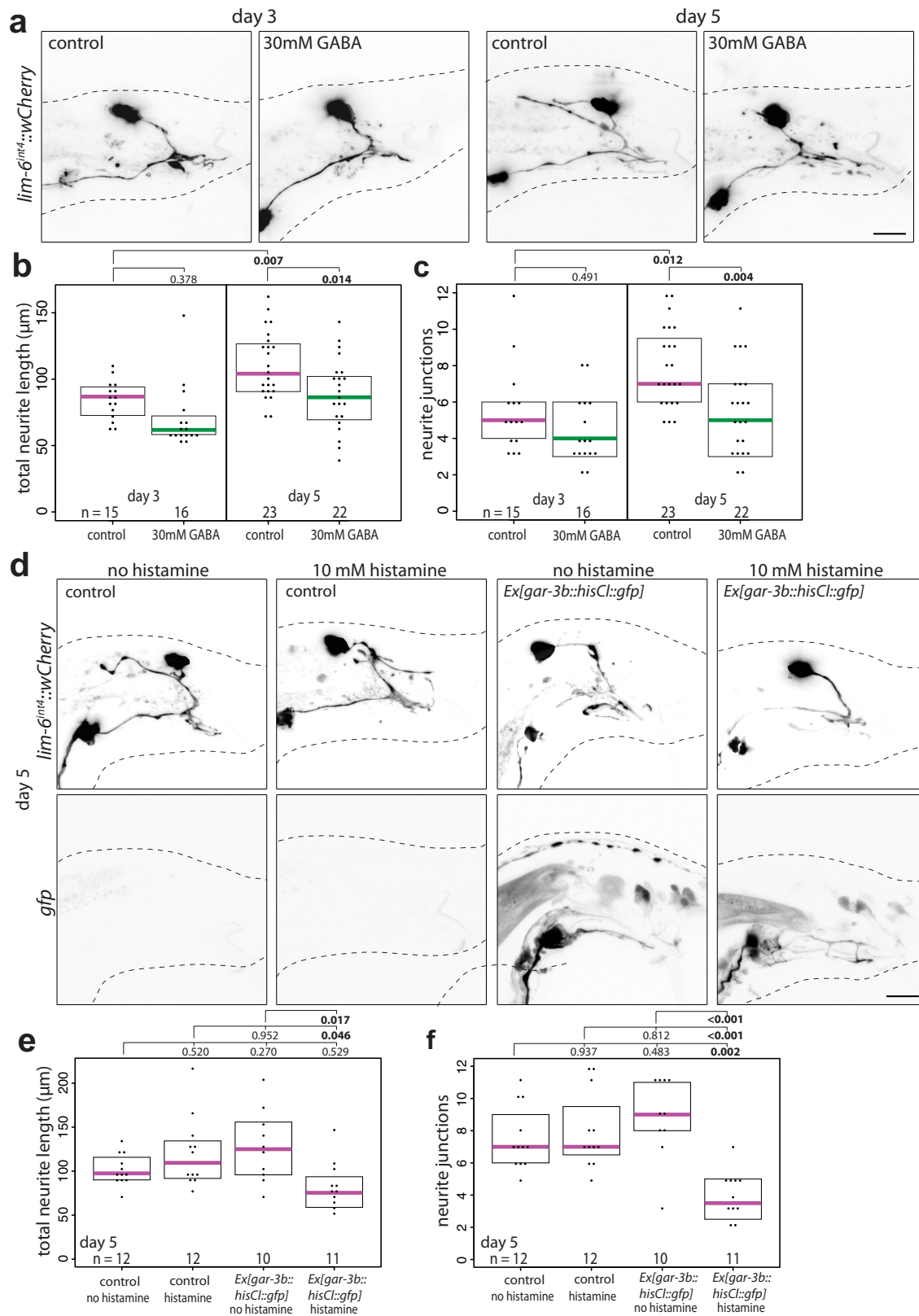
and differential interference contrast at day 1 in ventral and lateral views. Inset showing DVB and CP6 axons, with schematic of axons demonstrating lack of contact (red is DVB axon, green is CP6 axon, blue dashed lines are spicule retractor muscles). Asterisks in *flp-13::gfp* panel mark spicule retractor muscles. Dot represents one worm; magenta bar, median; boxes, quartiles; one-way ANOVA and post-hoc Tukey HSD, P values shown above plots, bold shows significance ($P < 0.05$), scale bars, 10 μ m.



Extended Data Figure 5 | See next page for caption.

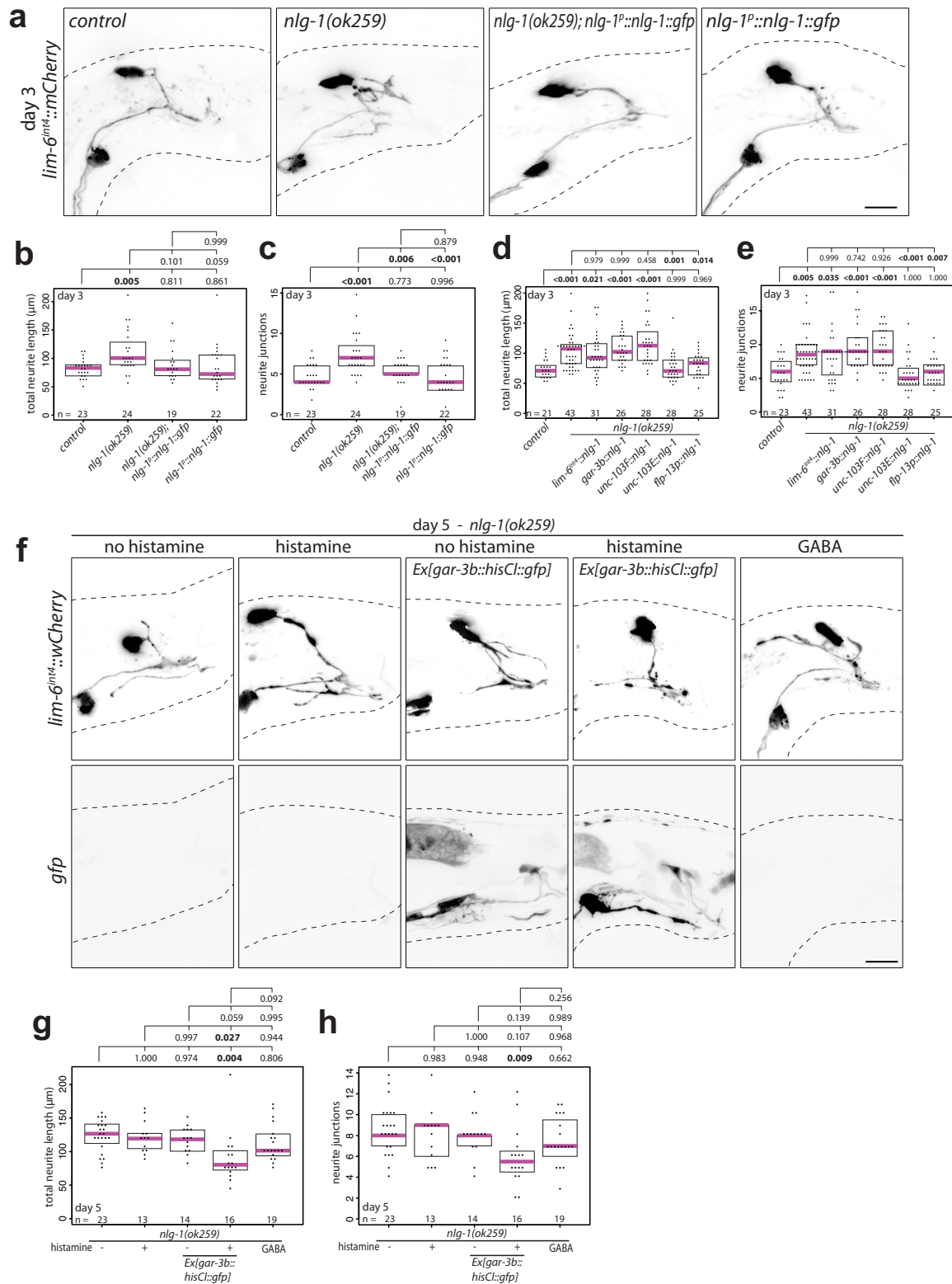
Extended Data Figure 5 | Day 1 male mating defects involving spicule coordination, spicule circuit activation in *unc-25*, *unc-97*, and *nrx-1* mutant males, and spicule neuron or muscle activation induces DVB neurite outgrowth. **a**, Per cent average mating success (sperm transfer) for day 1 and 3 males during 5-min timed mating assays with 15 *unc-31(e928)* hermaphrodites (*n* is number of worms, data points represent average percentage for each replicate of multiple males). **b**, Quantification of attempts at spicule prodding during 5-min timed mating assay for day 1 and 3 males. **c**, Ratio of protraction:prodding attempts during 5-min timed mating assay for males at days 1 and 3. **d–f**, Confocal images of *lim-6^{int4}::wCherry* (**d**), total neurite length (**e**), and number of neurite junctions (**f**) of *unc-25(e156)*, *unc-25(e156);Ex[gar-3b::ChR2::yfp]*, *unc-97(su110)*, *unc-97(su110);Ex[gar-3b::ChR2::yfp]*, *nrx-1(wy778)*, and *nrx-1(wy778);Ex[gar-3b::ChR2::yfp]* males following

activation at day 1 (488-nm light for 3×15 s every 45 min for 4.5 h). **g–i**, Confocal images (**g**) and quantification of total neurite outgrowth (**h**) and number of neurite junctions (**i**) in control, *Ex[unc-103E::ChR2::yfp]*, and *Ex[unc-103F⁴::ChR2::yfp]* worms after activation at day 1 with retinal (488-nm light for 3×5 s every 45 min for 4.5 h). **j, k**, Quantification of total neurite outgrowth (**j**) and number of neurite junctions (**k**) at day 1 in control, *Ex[lim-6^{int4}::ChR2::yfp]* (DVB), *Ex[unc-103E::ChR2::yfp]* (neuron-specific), and *Ex[unc-103F⁴::ChR2::yfp]* (muscle-specific) males after activation but in the absence of retinal. **l**, Time to protraction of control and *Ex[lim-6^{int4}::ChR2::yfp]* males after day 1 activation in the absence of retinal. Dot represents one worm; magenta bar, median; boxes, quartiles; one-way ANOVA and post-hoc Tukey HSD, *P* values shown above plots, bold shows significance ($P < 0.05$), scale bars, 10 μ m.



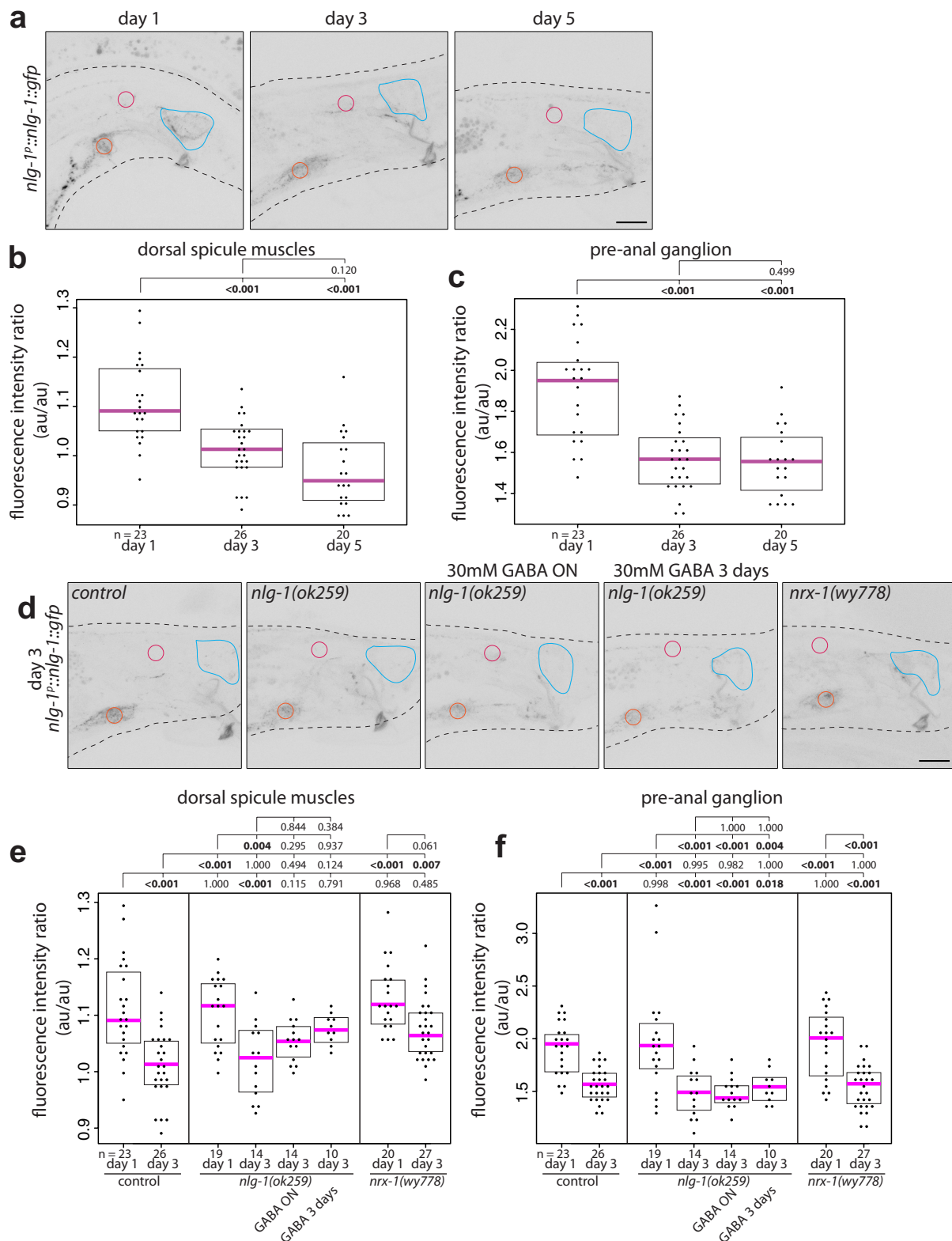
Extended Data Figure 6 | Exposure to exogenous GABA or silencing of spicule protraction circuit activity overnight reduces DVB neurites on day 5. **a–c**, Confocal images of *lim-6^{int4}::wCherry* (a), total neurite length (b), and number of neurite junctions (c) of males exposed overnight to 30 mM GABA at days 3 and 5. **d–f**, Confocal images of *lim-6^{int4}::wCherry* (d), total neurite length (e), and number of neurite junctions (f) at day 5 of

control worms with or without overnight 10 mM histamine, and *gar-3b::HisCl1::gfp* worms with or without overnight 10 mM histamine. Dot represents one worm; magenta bar, median; boxes, quartiles; one-way ANOVA and post-hoc Tukey HSD, *P* values shown above plots, bold shows significance (*P* < 0.05), scale bars, 10 μ m.



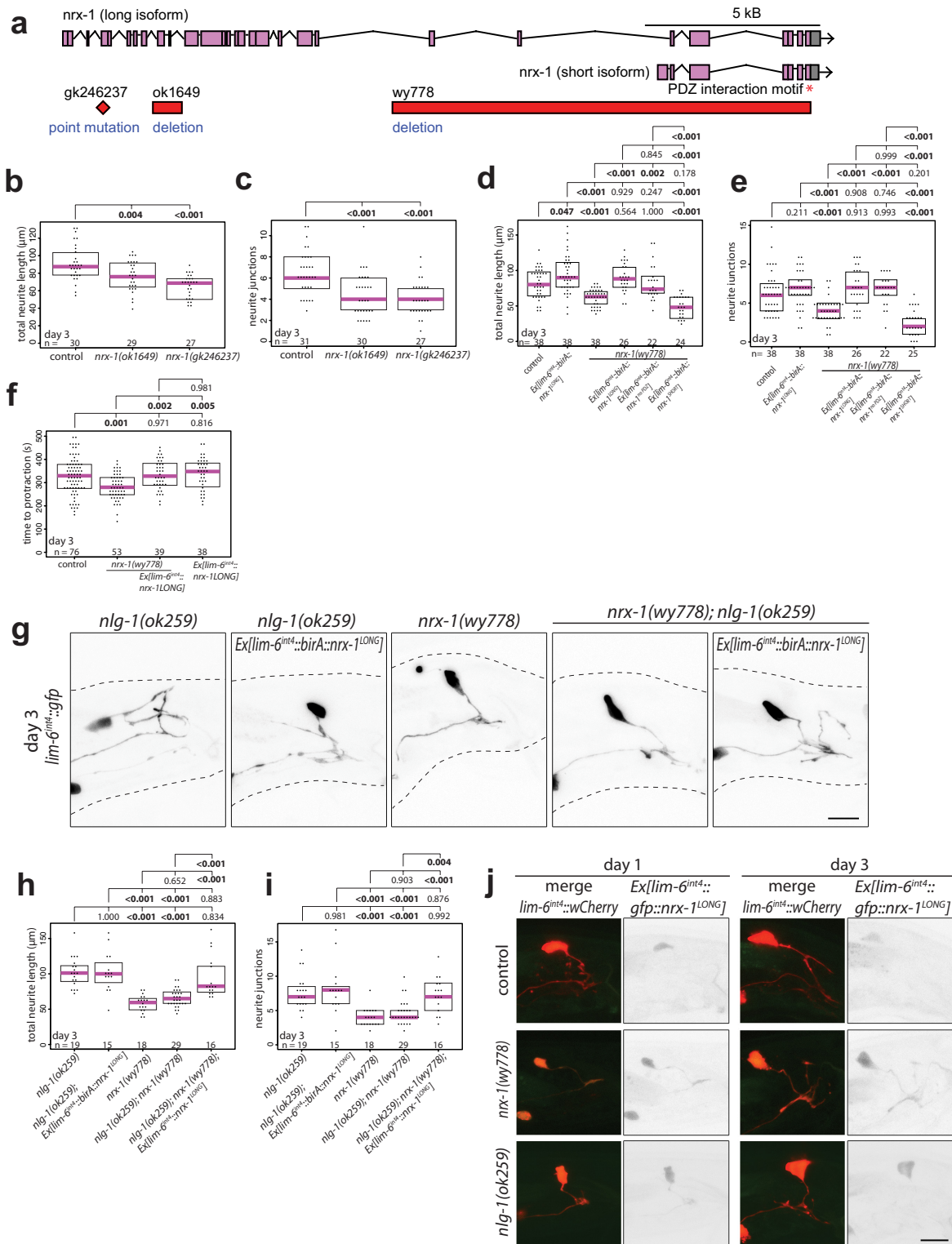
Extended Data Figure 7 | NLG-1 expression in multiple male sex muscles rescues *nlg-1* mutant DVB neurite phenotype. Silencing spicule circuit or exposure to exogenous GABA does not reduce DVB neurites in *nlg-1* mutant males. **a–c**, Confocal images of DVB (*lim-6^{int4}::wCherry*) (**a**), and quantification of total neurite outgrowth (**b**) and number of neurite junctions (**c**) in control, *nlg-1(ok259)*, *nlg-1(ok259); nlg-1^P::nlg-1:gfp*, and *nlg-1^P::nlg-1:gfp* day 3 males. **d, e**, Quantification of total neurite outgrowth (**d**) and number of neurite junctions (**e**) in control or *nlg-1(ok259)* mutant males with or without tissue-specific NLG-1 expression. Expression patterns for rescue promoters: *lim-6^{int4}* in DVB; *gar-3b* in SPC and spicule protractor muscles; *unc-103F* in SPC, PCA,

PCB and other neurons; *unc-103E* in male sex muscles; *flp-13* in spicule retractor muscles and CP6. **f–h**, Confocal images (**f**) of *lim-6^{int4}::wCherry* and *Ex[gar-3b::HisCl1::gfp]* in day 5 male worms, with total neurite length (**g**) and number of neurite junctions (**h**) of *nlg-1(ok259)* worms with or without 10 mM histamine overnight, *nlg-1(ok259); gar-3b::HisCl1::gfp* worms with or without 10 mM histamine overnight, and *nlg-1(ok259)* worms with 30 mM GABA overnight. Dot represents one worm; magenta bar, median; boxes, quartiles; one-way ANOVA and post-hoc Tukey HSD, *P* values shown above plots, bold shows significance (*P* < 0.05), scale bars, 10 μm.



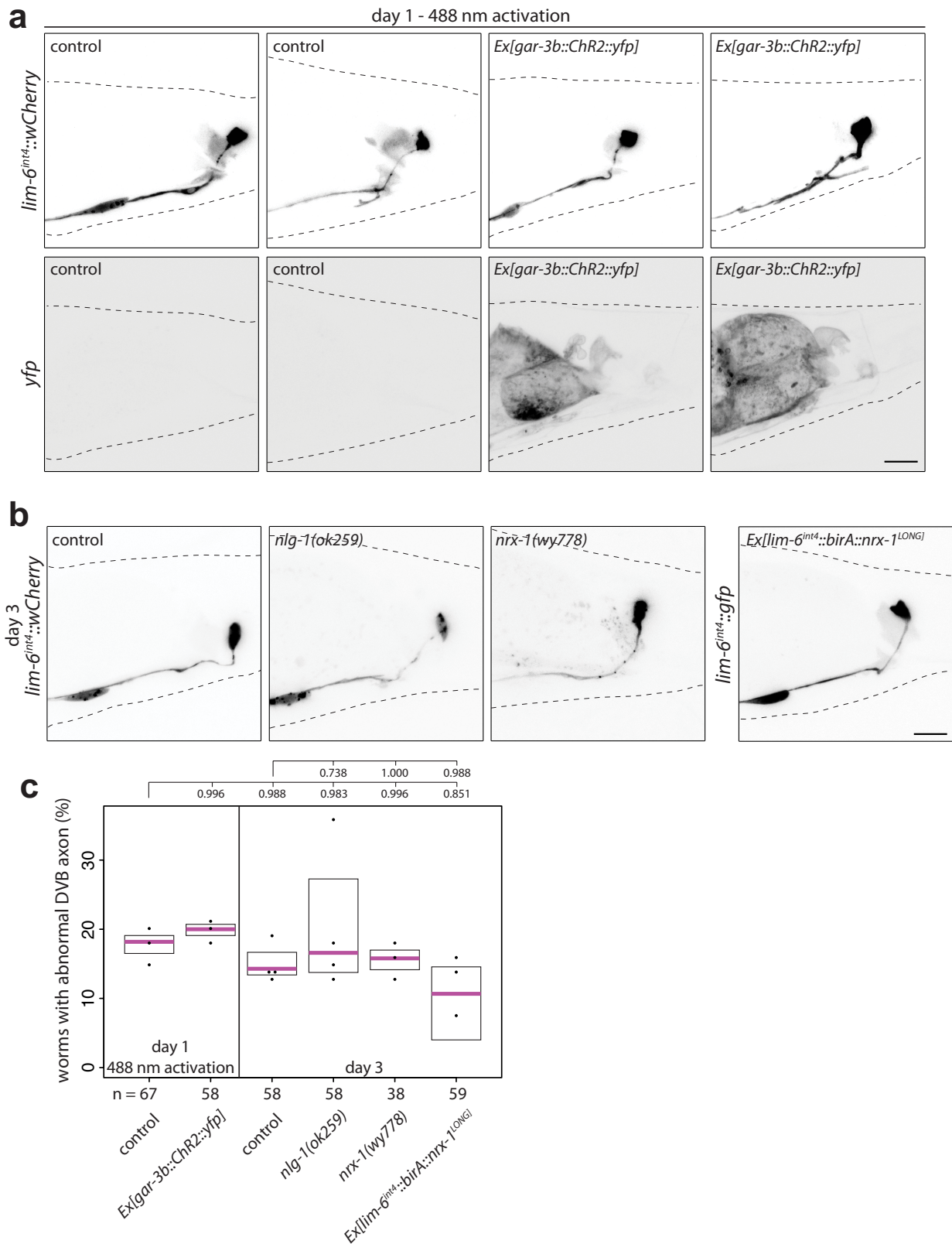
Extended Data Figure 8 | NLG-1 expression decreases from day 1 to day 3. **a**, Confocal images of *nlg-1p::nlg-1::gfp* in males at days 1, 3, and 5. Example regions of interest for measurements taken from single planes: blue, dorsal spicule muscles; red, pre-anal ganglion; magenta, DVB. **b**, **c**, Quantification of fluorescence intensity of *nlg-1p::nlg-1::gfp* in males at days 1, 3, and 5 reported as a ratio of mean fluorescence in dorsal spicule muscles (**b**) or pre-anal ganglion (**c**) normalized to background of DVB, which has little to undetectable expression. Dorsal spicule muscles refer to the gubernacular retractor, gubernacular erector, anterior oblique, and anal depressor. **d**, Confocal images of *nlg-1p::nlg-1::gfp* in day 3 males

as follows: control, *nlg-1(ok259)*, *nlg-1(ok259)* with overnight GABA exposure, *nlg-1(ok259)* with 3-day GABA exposure, and *nxr-1(wy778)*. **e**, **f**, Quantification of fluorescence intensity of *nlg-1p::nlg-1::gfp* in day 1 and 3 male worms as follows: control, *nlg-1(ok259)*, *nxr-1(wy778)*, day 3 *nlg-1(ok259)* with overnight GABA exposure, and *nlg-1(ok259)* with 3-day GABA exposure, as a ratio of mean fluorescence in dorsal spicule muscles (**e**) or pre-anal ganglion (**f**) normalized to background of DVB. Dot represents one worm; magenta bar, median; boxes, quartiles; one-way ANOVA and post-hoc Tukey HSD, *P* values shown above plots, bold shows significance (*P* < 0.05), scale bars, 10 μ m.



Extended Data Figure 9 | NRX-1 long isoform functions in DVB to control DVB neurite outgrowth and NRX-1 expression in DVB controls neurite outgrowth of *nlg-1* mutants. a, Genetic loci of *nrx-1* showing long and short isoforms, PDZ binding motif, and locations of point mutation gk246237 and deletions ok1649 and wy778. **b, c**, Quantification of total neurite length (**b**) and number of neurite junctions (**c**) in controls and long-isoform-specific mutants *nrx-1(ok1649)* and *nrx-1(gk246237)* at day 3. **d, e**, Quantification of total neurite outgrowth (**d**) and number of neurite junctions (**e**) at day 3 in control, *Ex[lim-6^{int4}::birA::nrx-1^{LONG}]*, *nrx-1(wy778)*, *nrx-1(wy778); Ex[lim-6^{int4}::birA::nrx-1^{LONG}]*, *nrx-1(wy778); Ex[lim-6^{int4}::birA::nrx-1^{SHORT}]*, and *nrx-1(wy778); Ex[lim-6^{int4}::birA::nrx-1^{noPDZ}]* worms. **f**, Time to spicule protraction

at day 3 in control, *nrx-1(wy778)*, *nrx-1(wy778); Ex[lim-6^{int4}::birA::nrx-1^{LONG}]*, and *Ex[lim-6^{int4}::birA::nrx-1^{LONG}]* worms. **g–i**, Confocal images of *lim-6^{int4}::wCherry* expression (**g**) and quantification of total neurite length (**h**) and number of neurite junctions (**i**) of day 3 *nlg-1(ok259)*, *nlg-1(ok259); Ex[lim-6^{int4}::birA::nrx-1^{LONG}]*, *nrx-1(wy778)*, *nlg-1(ok259); nrx-1(wy778)*, and *nrx-1(wy778); nlg-1(ok259); Ex[lim-6^{int4}::birA::nrx-1^{LONG}]* males. **j**, Confocal images of *lim-6^{int4}::wCherry* and *Ex[lim-6^{int4}::gfp::nrx-1^{LONG}]* in control, *nrx-1(wy778)*, and *nlg-1(ok259)* males at day 1 and 3. Dot represents one worm; magenta bar, median; boxes, quartiles; one-way ANOVA and post-hoc Tukey HSD, *P* values shown above plots, bold shows significance (*P* < 0.05) scale bars, 10 μm.



Extended Data Figure 10 | DVB in hermaphrodites does not show neurite branching upon *gar-3b::ChR2::yfp* activation or NRX-1 or NLG-1 manipulation. **a**, Confocal images of *lim-6^{int4}::wCherry* and *Ex[gar-3b::ChR2::yfp]* expression in day 1 hermaphrodites showing DVB axon projection after activation with retinal (488-nm light for 3×15 s every 45 min for 4.5 h). **b**, Confocal images of *lim-6^{int4}::wCherry* or *lim-6^{int4}::gfp* in control, *nrx-1(wy778)*, *nlg-1(ok259)*, and *Ex[lim-6^{int4}::gfp::nrx-1^{LONG}]* hermaphrodites at day 3. **c**, Quantification of the percentage of hermaphrodites with DVB axon abnormalities or

neurites (in almost all cases, a single neurite off the axon just posterior to the pre-anal ganglion) in day 1 control and *Ex[gar-3b::ChR2::yfp]* with activation, day 3 control, *nrx-1(wy778)*, *nlg-1(ok259)*, and *Ex[lim-6^{int4}::gfp::nrx-1^{LONG}]* worms. *n* shows number of worms, data points represent average percentage for each replicate of multiple hermaphrodites. Dot represents one worm; magenta bar, median; boxes, quartiles; one-way ANOVA and post-hoc Tukey HSD, *P* values shown above plots, bold shows significance ($P < 0.05$), scale bars, 10 μ m.

Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells

Juan I. Garaycoechea¹, Gerry P. Crossan¹, Frédéric Langevin¹, Lee Mulderrig¹, Sandra Louzada², Fentang Yang², Guillaume Guilbaud¹, Naomi Park², Sophie Roerink², Serena Nik-Zainal², Michael R. Stratton² & Ketan J. Patel^{1,3}

Haematopoietic stem cells renew blood. Accumulation of DNA damage in these cells promotes their decline, while misrepair of this damage initiates malignancies. Here we describe the features and mutational landscape of DNA damage caused by acetaldehyde, an endogenous and alcohol-derived metabolite. This damage results in DNA double-stranded breaks that, despite stimulating recombination repair, also cause chromosome rearrangements. We combined transplantation of single haematopoietic stem cells with whole-genome sequencing to show that this damage occurs in stem cells, leading to deletions and rearrangements that are indicative of microhomology-mediated end-joining repair. Moreover, deletion of p53 completely rescues the survival of aldehyde-stressed and mutated haematopoietic stem cells, but does not change the pattern or the intensity of genome instability within individual stem cells. These findings characterize the mutation of the stem-cell genome by an alcohol-derived and endogenous source of DNA damage. Furthermore, we identify how the choice of DNA-repair pathway and a stringent p53 response limit the transmission of aldehyde-induced mutations in stem cells.

The consumption of alcohol contributes to global mortality and cancer development¹. Most of the toxic effects of alcohol are probably caused by its oxidation product acetaldehyde, which is highly reactive towards DNA². The enzyme aldehyde dehydrogenase 2 (ALDH2) prevents acetaldehyde accumulation by oxidizing it efficiently to acetate, but around 540 million people carry a polymorphism in *ALDH2* that encodes a dominant-negative variant of the enzyme³. Alcohol consumption in these individuals induces an aversive reaction and predisposes them to oesophageal cancer⁴. Nevertheless, ALDH2 deficiency is surprisingly well tolerated in humans. This could be because of the additional tier of protection provided by FANCD2, a DNA-crosslink-repair protein. In fact, genetic inactivation of *Aldh2* and *Fancd2* in mice leads to cancer and a profound haematopoietic phenotype^{5,6}. In humans, deficiency in DNA-crosslink repair causes the inherited illness Fanconi anaemia, a devastating condition that leads to abnormal development, bone-marrow failure and cancer⁷. Acetaldehyde genotoxicity is likely to contribute to this phenotype, as Japanese children who are afflicted with Fanconi anaemia and carry the *ALDH2* polymorphism display earlier-onset bone marrow failure⁸. Together, these data suggest that endogenous aldehydes are a ubiquitous source of DNA damage that impairs blood production.

It is likely that some of this damage occurs in haematopoietic stem cells (HSCs), which are responsible for lifelong blood production. HSC attrition is a feature of ageing, and mutagenesis in the remaining HSCs promotes dysfunctional haematopoiesis and leukaemia. Moreover, both humans and mice that lack DNA repair factors are prone to HSC loss, and in some cases, bone marrow failure^{9,10}. HSCs employ DNA repair and respond to damage in a distinct manner compared to later progenitors^{11,12}. While these observations point to a fundamental role for DNA repair in HSCs, recent work has highlighted that the response to replication stress maintains HSC function and integrity¹³. However, there is a key gap in our knowledge regarding the identity of the endogenous factors that damage DNA and lead to replication stress. Here we show that alcohol-derived and endogenous aldehydes damage the genomes of haematopoietic cells, and we characterize the surveillance

and repair mechanisms that counteract this. We also establish a method that allows us to determine the mutational landscape of individual HSCs, and in doing so, provide new insight into the p53 response in mutagenized stem cells.

Ethanol stimulates homologous recombination repair

Aldh2^{-/-} *Fancd2*^{-/-} mice develop severe HSC attrition, causing spontaneous bone marrow failure, which can also be induced by exposing these mice to ethanol^{5,6}. This genetic interaction suggests that in the absence of aldehyde catabolism (such as in *Aldh2*^{-/-} mice), DNA repair is engaged to maintain blood homeostasis. To test this theory, we set out to monitor DNA repair activity *in vivo*. The Fanconi anaemia pathway repairs DNA crosslinks by using a replication-coupled excision mechanism that is completed by homologous recombination^{14,15}. We therefore used a method to visualize sister-chromatid exchange (SCE) events in bone marrow cells of living mice; these represent recombination repair transactions coupled to replication (Fig. 1a). The number of SCE events is elevated 2.3-fold in *Aldh2*^{-/-} mice, indicating that recombination repair is stimulated in response to endogenous aldehydes (Fig. 1b, c). Moreover, a single exposure to alcohol causes a fourfold increase in SCE events in *Aldh2*^{-/-} mice (Fig. 1b, c, Extended Data Fig. 1a), suggesting that physiological acetaldehyde accumulation in blood cells is not sufficient to inactivate the homologous recombination repair factor BRCA2¹⁶. *Fancd2*^{-/-} mice do not show a similar induction following exposure to ethanol; therefore, detoxification is the primary mechanism that prevents DNA damage by aldehydes and alcohol. Finally, the number of SCE events in *Aldh2*^{-/-} *Fancd2*^{-/-} mice is indistinguishable from that in *Aldh2*^{-/-} mice, showing that homologous recombination repair occurs despite inactivation of FANCD2 (Fig. 1c, Extended Data Fig. 1b).

The repair of aldehyde-induced DNA damage is therefore not limited to the Fanconi anaemia crosslink-repair pathway. As the recombination machinery is essential for mouse development, we used the isogenic chicken B-cell line DT40, which has been used to define the involvement of homologous recombination in crosslink repair¹⁴.

¹MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Avenue, Cambridge CB2 0QH, UK. ²Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK.

³Department of Medicine, University of Cambridge, Addenbrooke's Hospital, Hills Rd, Cambridge CB2 0QQ, UK.

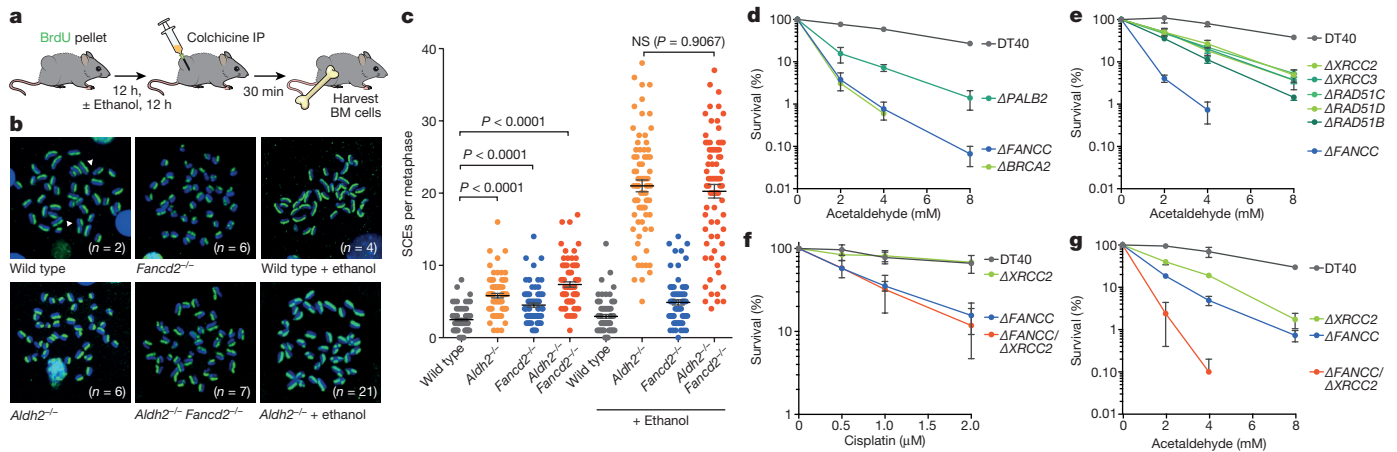


Figure 1 | Ethanol induces potent homologous recombination *in vivo*. **a**, Treatment of mice with BrdU for differential labelling of sister chromatids of bone marrow cells *in vivo*. Some mice were also treated with ethanol, a precursor of acetaldehyde. IP, intraperitoneal injection; BM, bone marrow. **b**, Representative images of bone marrow metaphase spreads (*n*, number of SCEs per metaphase). **c**, Number of SCEs in

the bone marrow of *Aldh2*^{-/-} *Fancd2*^{-/-} and control mice (triplicate experiments, 25 metaphases per mouse, *n* = 75; *P* calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.). NS, not significant. **d–g**, Clonogenic survival of DT40 DNA-repair mutants (triplicate experiments; data shown as mean and s.e.m.).

DT40 cells carrying disruptions of key homologous recombination genes show hypersensitivity to acetaldehyde (Fig. 1d, e), in a similar way to cells lacking the Fanconi anaemia gene *FANCC*. To test the relationship between the Fanconi anaemia and homologous recombination pathways, we analysed the sensitivity of cells deficient in both *FANCC* and *XRCC2*. These cells showed the same sensitivity to cisplatin as the single knockout cells (Fig. 1f), but were much more sensitive to acetaldehyde (Fig. 1g), indicating that homologous recombination repair confers additional acetaldehyde resistance beyond that provided by Fanconi anaemia crosslink repair. In summary, detoxification provides the dominant protection mechanism against endogenous aldehydes; however, when aldehydes damage DNA, cells use both DNA-crosslink and homologous recombination repair.

FANCD2 prevents alcohol-induced genomic instability

The active DNA recombination in bone marrow cells indicates that even in the absence of *FANCD2*, there is an alternative repair response to both endogenous and ethanol-derived aldehydes. However, our previous work has shown that *Aldh2*^{-/-} *Fancd2*^{-/-} mice lose the ability to maintain blood production^{5,6}. To determine whether this is due to the accumulation of damaged DNA, we examined haematopoietic cells for evidence of broken chromosomes. One marker of genetic instability is the formation of micronuclei, which are formed from lagging or broken chromosomes. Micronuclei are easily quantified in normochromic erythrocytes (NCEs) *in vivo*, because they persist following enucleation (Fig. 2a, Extended Data Fig. 1c). There is a significant increase in the proportion of NCEs with micronuclei in

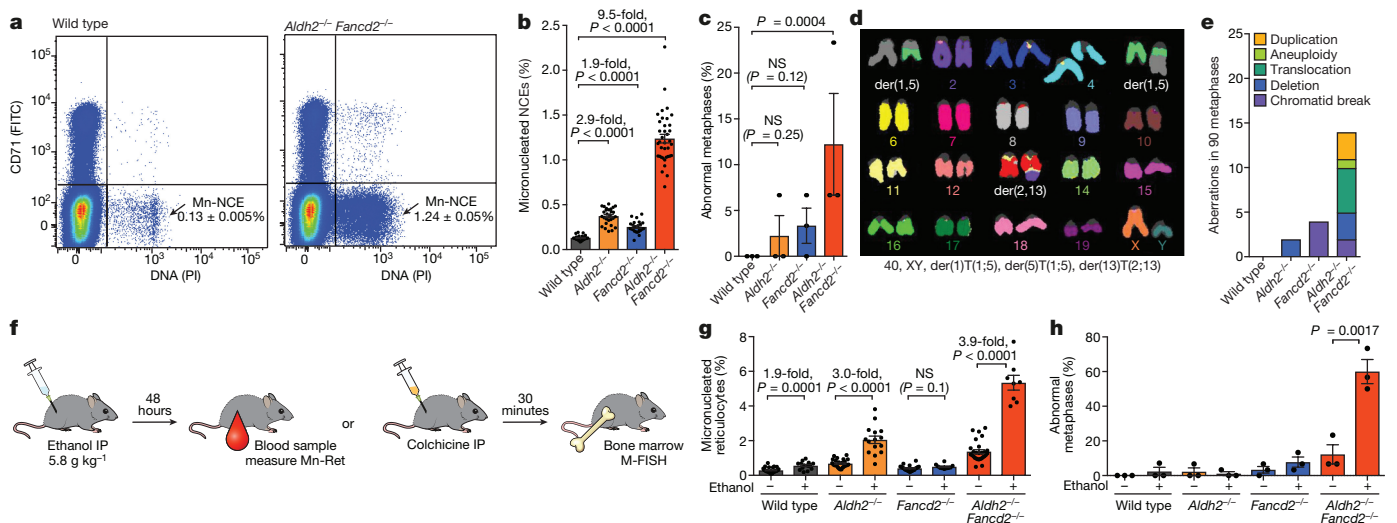


Figure 2 | Spontaneous and ethanol-induced genomic instability in *Aldh2*^{-/-} *Fancd2*^{-/-} mice. **a**, Quantification of micronucleated normochromic erythrocytes (Mn-NCE, CD71⁺ PI⁺) by flow cytometry. **b**, Percentage of micronucleated normochromic erythrocytes (*P* calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; *n* = 28, 28, 25 and 37 mice, left to right). **c**, Percentage of abnormal metaphases in bone marrow cells (*P* calculated by one-sided Fisher's exact test; data shown as mean and s.e.m.; three mice per genotype, 30 metaphases per mouse). **d**, A *Aldh2*^{-/-} *Fancd2*^{-/-} metaphase, showing two translocations, see Extended Data Fig. 1f–i for the complete list of aberrations. **e**, Types of

chromosomal aberrations (90 metaphases per genotype). **f**, Treatment of mice with ethanol to assess genomic instability with the micronucleus assay (**g**) or M-FISH karyotyping (**h**). **g**, Percentage of micronucleated reticulocytes (Mn-Ret, CD71⁺ PI⁺) after ethanol treatment. *P* calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; *n* = 29, 15, 25, 15, 20, 10, 28 and 9 mice, left to right. **h**, Abnormal metaphases in bone marrow cells after ethanol treatment. *P* calculated by one-sided Fisher's exact test; data shown as mean and s.e.m.; 3 mice per genotype, 30 metaphases per mouse.

both *Aldh2*^{-/-} (2.9-fold) and *Fancd2*^{-/-} (1.9-fold) mice compared to wild-type controls, but the increase is much larger in *Aldh2*^{-/-} *Fancd2*^{-/-} mice (9.5-fold, Fig. 2b). These micronuclei could represent genomic instability during blood production. We therefore examined cells in metaphase obtained directly from the bone marrow of these mice with multiplex fluorescence *in situ* hybridization (M-FISH). More than 10% of *Aldh2*^{-/-} *Fancd2*^{-/-} bone-marrow cells carried chromosomal aberrations, encompassing all classes of cytogenetic change (Fig. 2c–e). These aberrations are not clonal events, because each karyotype was unique (Extended Data Fig. 1).

Next, we investigated whether this chromosome damage was exacerbated by exposure to ethanol. As a control, we exposed wild-type or *Fancd2*^{-/-} mice to mitomycin C (Extended Data Fig. 1d). The experimental scheme (outlined in Fig. 2f) shows how we determined the prevalence of micronuclei in reticulocytes and aberrant metaphases following exposure to ethanol. A single dose of ethanol caused a marked increase in the proportion of reticulocytes containing micronuclei in *Aldh2*^{-/-} mice. Notably, this induction was comparable to that observed in wild-type mice following exposure to agents known to induce genome instability, such as ionizing irradiation or vincristine (Extended Data Fig. 1e). However, there was a stronger induction of micronucleus formation in *Aldh2*^{-/-} *Fancd2*^{-/-} mice than in controls (Fig. 2g), which was accompanied by a striking increase in the number of abnormal metaphases, with almost 60% of metaphases having damaged chromosomes following ethanol exposure (Fig. 2h, Extended Data Fig. 1g–i). These mice rapidly lost the ability to produce blood and died from bone-marrow failure (Extended Data Fig. 2). These results show that, despite activation of homologous recombination, the Fanconi anaemia crosslink-repair pathway is essential for preventing chromosome breakage and loss of blood homeostasis in response to aldehydes.

Ku70 contributes to repair of aldehyde-induced DSBs

The presence of chromosome breaks and translocations suggests that aldehydes cause double-stranded breaks (DSBs), which could be processed by non-homologous end-joining (NHEJ) repair¹⁷. Previous studies in cell lines and nematodes have indicated that, in the absence of the Fanconi anaemia pathway, engagement of DSBs by NHEJ leads to further genomic instability^{18,19}. Therefore, we investigated whether NHEJ and Fanconi anaemia repair are redundant in resolving endogenous DNA damage in HSCs, and whether there is a role for NHEJ in maintaining resistance to acetaldehyde.

To do this, we crossed mice deficient in the known Fanconi anaemia repair gene *Fanca* with mice lacking the key NHEJ factor Ku70 (encoded by *Xrcc6*, also known as *Ku70*). We failed to obtain *Fanca*^{-/-} *Ku70*^{-/-} mice, indicating that there was a synthetic lethal interaction between Ku70-dependent NHEJ and Fanconi anaemia crosslink repair (Supplementary Information Table 1). To bypass embryonic lethality, we generated blood-specific *Fanca* knock-out mice (Extended Data Fig. 3) and crossed them with *Ku70*^{+/-} mice to produce mice that had the double mutation in HSCs and the blood compartment (*Fanca*^{fl/-} *Ku70*^{-/-} *Vav1*-iCre). These mice were viable, indicating that the embryonic lethality of *Fanca*^{-/-} *Ku70*^{-/-} is not due to failed blood production (Supplementary Information Table 1). However, blood counts show that *Fanca*^{fl/-} *Ku70*^{-/-} *Vav1*-iCre mice are anaemic (Fig. 3a) and have fewer HSCs compared to congenic controls (Fig. 3b–e). *Fanca*^{fl/-} *Ku70*^{-/-} *Vav1*-iCre mice also display genomic instability, with increased frequency of micronuclei-containing NCEs (Fig. 3f). Finally, we tested whether Ku70 was required to maintain resistance of short-term (ST)-HSCs to aldehydes by exposing bone marrow cells to acetaldehyde *in vitro* before injecting them into lethally irradiated recipients. *Fanca*^{fl/-} *Ku70*^{-/-} *Vav1*-iCre ST-HSCs were much more sensitive to acetaldehyde than either of the single mutant ST-HSCs (Fig. 3g). These results indicate that in the mouse haematopoietic system, in the absence of Fanconi anaemia repair, NHEJ is required to provide resistance to endogenous and acetaldehyde-induced DNA

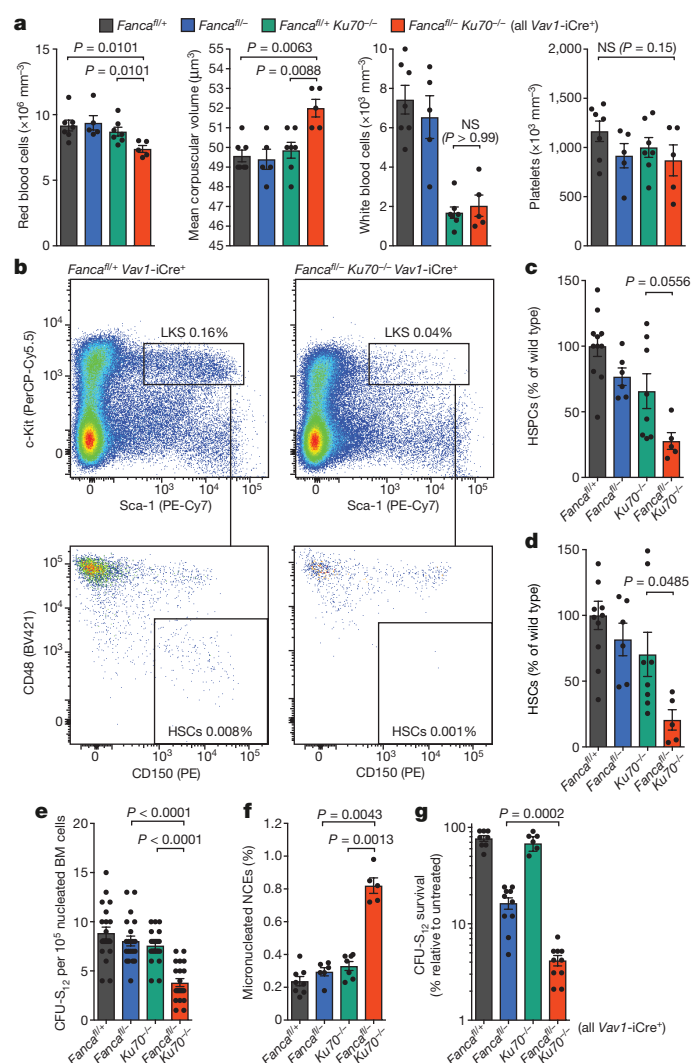


Figure 3 | NHEJ cooperates with the Fanconi anaemia pathway to maintain HSC integrity, genomic stability and cellular resistance to aldehydes. **a**, Blood parameters of 8- to 12-week old mice (P calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; $n = 8, 6, 7$ and 5 mice, left to right). **b**, Representative flow cytometry plot of haematopoietic stem and progenitor cells (HSPCs) of *Fanca*^{fl/-} *Ku70*^{-/-} *Vav1*-iCre mice and control. LKS, Lin⁻Kit⁺Sca-1⁺. **c**, **d**, Quantification of HSPCs (Lin⁻Kit⁺Sca-1⁺) and HSCs (Lin⁻Kit⁺Sca-1⁺CD48⁺CD150⁺) by flow cytometry (P calculated by two-tailed Student's t -test; data shown as mean and s.e.m.; n as in **a**). **e**, Counts of colony-forming unit-spleen (CFU-S) colonies from the bone marrow of *Fanca*^{fl/-} *Ku70*^{-/-} *Vav1*-iCre and control mice. Each point represents the number of CFU-S₁₂ in a single recipient (P calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; $n = 20$ mice). **f**, Frequency of Mn-NCE (P calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; n as in **a**). **g**, Survival of CFU-S₁₂ after treatment with 4 mM acetaldehyde for 4 h, relative to untreated samples (P calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; $n = 10$ mice).

damage. This result contrasts with the reported negative impact of active NHEJ on the viability of Fanconi anaemia-deficient chicken DT40 cells and worms^{18,19}.

Aldehyde-damaged HSCs are functionally compromised

Our results so far indicate that endogenous aldehydes give rise to DSBs in the absence of Fanconi anaemia repair, which are engaged by homologous recombination and NHEJ, but ultimately rearrange chromosomes in bone marrow cells. A key question is whether endogenous DNA damage and subsequent mutations accumulate in the HSC

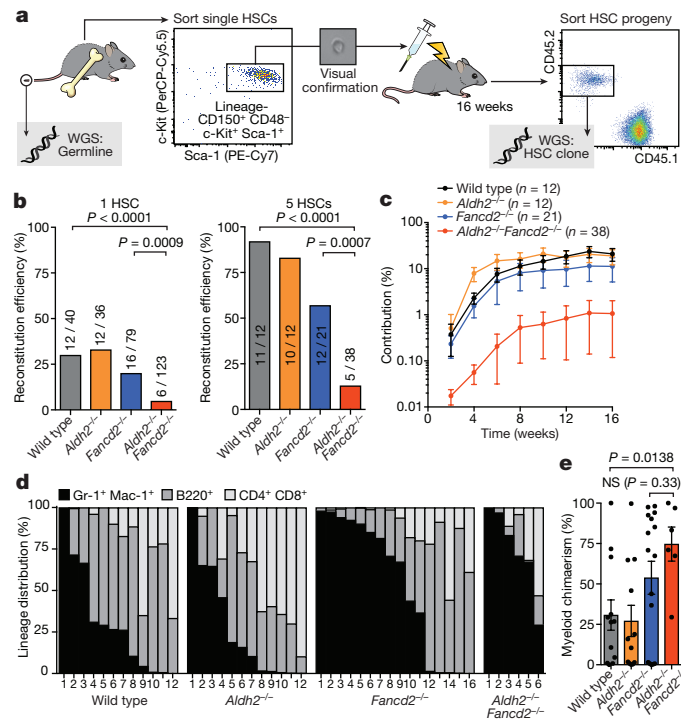


Figure 4 | Single HSC transplantation reveals that *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs are functionally compromised. **a**, Transplantation of single HSCs for the generation of HSC clones *in vivo*. The HSC progeny (CD45.2⁺) were recovered after four months and analysed by whole-genome sequencing, alongside a germline reference. **b**, Percentage and number of irradiated recipients that were positive for reconstitution by one or five transplanted HSCs (*P* calculated by two-sided Fisher's exact test).

compartment. This is a critical question because there is evidence that HSCs differ in their DNA repair capacity and response compared to later progenitors¹¹. Two obstacles had to be overcome in order to establish whether endogenous aldehydes mutate the genomes of these vital cells. First, the stochastic nature of DNA damage makes it unlikely that the same mutation will occur in multiple cells. Second, the scarcity of HSCs, especially in the case of *Aldh2*^{-/-} *Fancd2*^{-/-} mice, precludes the use of most conventional techniques to assess DNA damage. We also wanted to ascertain whether mutations arise in functional stem cells, and therefore avoided whole-genome amplification or short-term *in vitro* expansion of cells isolated by flow cytometry. Instead, we decided to define HSCs functionally and exploit the ability of a single HSC to reconstitute long-term blood production following transplantation into a lethally irradiated mouse²⁰.

Our approach combines transplantation of single HSCs with whole-genome sequencing to obtain the mutational landscape of stem cells, while also allowing us to assess the functional capacity of mutant HSCs (Fig. 4a). We carried out transplants with one or five *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs (Fig. 4b). These stem cells rarely engrafted (with a frequency of 4.8%), contributed less to haematopoiesis and were myeloid-biased compared to controls (Fig. 4b–e). These results indicate that *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs are severely functionally compromised and share features with aged HSCs^{21,22}.

Mutational landscape of aldehyde-damaged stem cells

Our ultimate goal was to obtain clonal blood, which provided us with a physiological method to amplify stem-cell genomes. As outlined (Fig. 4a), four months after transplantation, we isolated the CD45.2⁺ HSC progeny and performed whole-genome sequencing at 20× coverage; tail DNA from the donor mouse served as the germline reference. This allowed us to detect heterozygous somatic changes, which are absent in the matched germline reference and represent

mutations in the HSC. Genomes of *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs were mutated with increased prevalence of indels, rearrangements and translocations (Fig. 5a, Extended Data Fig. 4).

Although the number of single-base substitutions was significantly higher in *Aldh2*^{-/-} *Fancd2*^{-/-} genomes (Fig. 5b), the total numbers were low and no changes were detected in the type of substitutions (Fig. 5c). We also found no difference in the frequency or pattern of point mutations in bone marrow cells of *Aldh2*^{-/-} *Fancd2*^{-/-} mice using the Select-cII Big Blue *in vivo* reporter assay (Extended Data Fig. 5).

A limitation of our approach is that cells with the capacity to engraft may represent the least mutated HSCs. Nevertheless, we observed significant increases in the frequency of deletions, which were more prevalent (Fig. 5d, e) and larger (Fig. 5f) in *Aldh2*^{-/-} *Fancd2*^{-/-} genomes. The mean variant allele frequency (VAF) for all filtered indels was 0.47, establishing that these changes are of clonal origin. By examining the flanking regions, we found that microhomology-mediated deletions are the main contributors to the mutations observed in *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs, indicative of end-joining repair of DSBs²³ (Fig. 5g, h). Additionally, the increase in the size of the deletions (Fig. 5f) suggests a role for alternative end-joining in the repair of some of these breaks, as alternative end-joining is characterized by increased resection in comparison to classical NHEJ²⁴. Next, we analysed the location of indels across the genome, as recent work has suggested a role for the Fanconi anaemia pathway in preventing genomic instability at transcription–replication collisions^{25,26}. However, we found no evidence of microhomology-mediated deletions being enriched at coding regions or transcribed genes (Fig. 5i, j), suggesting that DSB formation in *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs is stochastic.

The most striking change in *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs was the presence of rearrangements that were not detected in most controls. *Aldh2*^{-/-} *Fancd2*^{-/-} stem cells contained on average two rearrangements per genome; in contrast, we observed only two large deletions among all ten control HSC genomes (Fig. 5k–l). In summary, these data

mutations in the HSC. Genomes of *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs were mutated with increased prevalence of indels, rearrangements and translocations (Fig. 5a, Extended Data Fig. 4).

Although the number of single-base substitutions was significantly higher in *Aldh2*^{-/-} *Fancd2*^{-/-} genomes (Fig. 5b), the total numbers were low and no changes were detected in the type of substitutions (Fig. 5c). We also found no difference in the frequency or pattern of point mutations in bone marrow cells of *Aldh2*^{-/-} *Fancd2*^{-/-} mice using the Select-cII Big Blue *in vivo* reporter assay (Extended Data Fig. 5).

A limitation of our approach is that cells with the capacity to engraft may represent the least mutated HSCs. Nevertheless, we observed significant increases in the frequency of deletions, which were more prevalent (Fig. 5d, e) and larger (Fig. 5f) in *Aldh2*^{-/-} *Fancd2*^{-/-} genomes. The mean variant allele frequency (VAF) for all filtered indels was 0.47, establishing that these changes are of clonal origin. By examining the flanking regions, we found that microhomology-mediated deletions are the main contributors to the mutations observed in *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs, indicative of end-joining repair of DSBs²³ (Fig. 5g, h). Additionally, the increase in the size of the deletions (Fig. 5f) suggests a role for alternative end-joining in the repair of some of these breaks, as alternative end-joining is characterized by increased resection in comparison to classical NHEJ²⁴. Next, we analysed the location of indels across the genome, as recent work has suggested a role for the Fanconi anaemia pathway in preventing genomic instability at transcription–replication collisions^{25,26}. However, we found no evidence of microhomology-mediated deletions being enriched at coding regions or transcribed genes (Fig. 5i, j), suggesting that DSB formation in *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs is stochastic.

The most striking change in *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs was the presence of rearrangements that were not detected in most controls. *Aldh2*^{-/-} *Fancd2*^{-/-} stem cells contained on average two rearrangements per genome; in contrast, we observed only two large deletions among all ten control HSC genomes (Fig. 5k–l). In summary, these data

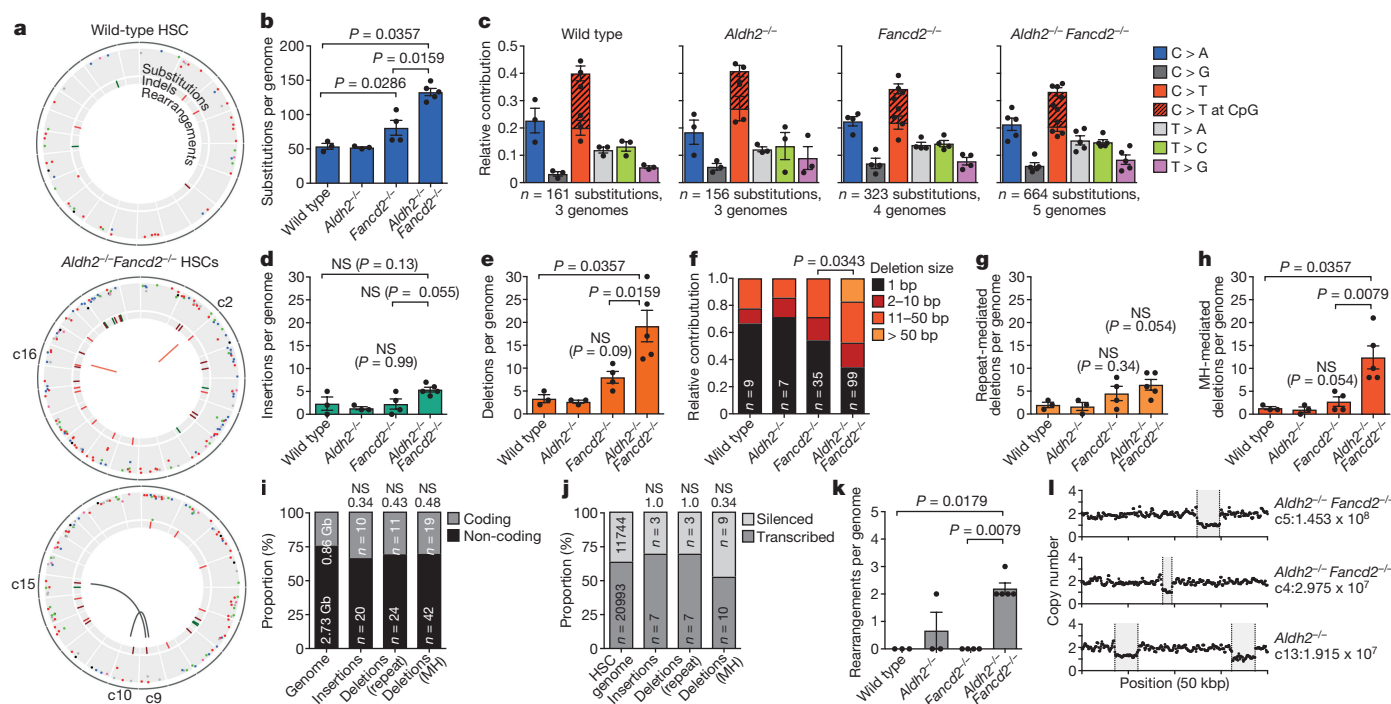


Figure 5 | Endogenous aldehydes mutate the HSC genome. **a**, Circos plots showing mutations in three HSCs. All HSCs are shown in Extended Data Fig. 4. **b**, **d**, **e**, **g**, **h**, **k**, Mutations of different classes per genome (*P* calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; *n* = 3, 3, 4 and 5 HSC genomes, left to right). **b**, Number of substitutions. **c**, Point mutation classes in HSC genotypes. **d**, Number of insertions per genome. **e**, Number of deletions per genome. **f**, Distribution of the size of deletions (χ^2 test, *n* shows number of deletions). **g**, Number

of repeat-mediated deletions per genome. **h**, Number of microhomology (MH)-mediated deletions per genome. **i**, **j**, Indels in *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs are randomly distributed: within or outside genes (i) (*P* calculated by hypergeometric distribution, *n* is number of indels), or between expressed or silenced genes (j) (*P* calculated by binomial distribution, *n* is number of indels). Numbers above columns, *P* values. **k**, Number of rearrangements per genome. **l**, Large copy-number losses in *Aldh2*^{-/-} *Fancd2*^{-/-} and *Aldh2*^{-/-} HSCs at the indicated locations.

provide the first whole-genome sequences obtained from single stem cells propagated *in vivo*. These stem cell genomes show that endogenous aldehydes induce a tapestry of inter-chromosomal changes that are mediated by mutagenic end-joining of DNA DSBs.

A p53 response removes aldehyde-damaged HSCs

Strikingly, most *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs failed to engraft (Fig. 4b). It is possible that HSCs that carry heavy DNA-damage burdens are eliminated, and selection pressure favours the survival of less damaged stem cells. It was therefore important to determine the mechanism of HSC loss in *Aldh2*^{-/-} *Fancd2*^{-/-} mice, and if attenuated, it would be important to determine the mutagenic consequences. The p53 protein regulates the cellular response to DNA damage and, when activated, induces restorative processes or apoptosis. We found that *Aldh2*^{-/-} *Fancd2*^{-/-} haematopoietic stem and progenitor cells (HSPCs) accumulated p53 and cleaved caspase-3, indicating that endogenous-aldehyde stress activates the p53 response (Extended Data Fig. 6a–c). Furthermore, we found that genetic ablation of p53 partially suppressed the acetaldehyde hypersensitivity of *Fancd2*-deficient splenic B cells and granulocyte/macrophage colony forming units (Extended Data Fig. 6d, e).

We therefore generated *Aldh2*^{-/-} *Fancd2*^{-/-} *Trp53*^{-/-} triple-knock-out mice. The severe HSC depletion of *Aldh2*^{-/-} *Fancd2*^{-/-} mice was completely rescued in the triple knockouts (Fig. 6a, b). In addition, the triple-knockout stem cells were functional, as the mice showed a complete rescue in the frequency of ST-HSCs upon bone-marrow transplantation (Extended Data Fig. 6f). Moreover, p53 deficiency fully restored the blood cytopenias of untreated *Aldh2*^{-/-} *Fancd2*^{-/-} mice and made these mice more resistant to alcohol exposure (Extended Data Fig. 7). Notably, *Trp53* deletion did not rescue the embryonic lethality of *Aldh2*^{-/-} *Fancd2*^{-/-} embryos (in *Aldh2*^{-/-} mothers), suggesting that a different checkpoint might mediate developmental failure (Supplementary Information Table 2).

We reasoned that the rescue of haematopoiesis in *Aldh2*^{-/-} *Fancd2*^{-/-} *Trp53*^{-/-} mice must be occurring at the cost of genome integrity. Although the level of micronucleated NCEs in the blood of *Aldh2*^{-/-} *Fancd2*^{-/-} *Trp53*^{-/-} mice appeared similar to that of *Aldh2*^{-/-} *Fancd2*^{-/-} mice (Extended Data Fig. 8a), we noticed a significant (*P* = 0.0034) increase in chromosome rearrangements in *Aldh2*^{-/-} *Fancd2*^{-/-} *Trp53*^{-/-} mice, as seen by M-FISH analysis of total bone marrow cells (Fig. 6c, Extended Data Fig. 8). However, neither of these analyses tell us whether genome stability is similarly compromised in *Aldh2*^{-/-} *Fancd2*^{-/-} *Trp53*^{-/-} HSCs. We therefore performed transplantation of single HSCs combined with whole-genome sequencing, as described earlier, and observed that p53 deficiency partially rescued the engraftment defect of *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs (Fig. 6d). Surprisingly, the genomes of *Aldh2*^{-/-} *Fancd2*^{-/-} *Trp53*^{-/-} stem cells did not carry a greater mutation burden compared to those of *Aldh2*^{-/-} *Fancd2*^{-/-} HSCs (Fig. 6e, f). Indel and rearrangement calls were validated by targeted deep sequencing and PCR, respectively (Extended Data Figs 9, 10 and Methods). One plausible explanation for the lack of increased mutagenesis in *Aldh2*^{-/-} *Fancd2*^{-/-} *Trp53*^{-/-} HSCs is the possibility that the very small number of HSCs in *Aldh2*^{-/-} *Fancd2*^{-/-} mice might have undergone more replicative cycles, thereby accruing a larger number of mutations. To address this, we quantified the frequency of ‘clock’ mutations (C to T at CpG sites), but this analysis showed no significant difference between *Aldh2*^{-/-} *Fancd2*^{-/-} and *Aldh2*^{-/-} *Fancd2*^{-/-} *Trp53*^{-/-} HSCs (Fig. 6f). These results indicate that aldehyde-induced DNA damage induces p53 leading to HSC attrition, which is inconsistent with p53 being a negative regulator of Fanconi anaemia repair, as recently reported²⁷. However, while *Trp53* deletion completely rescues HSC depletion, this does not occur at the expense of genome stability in blood stem cells.

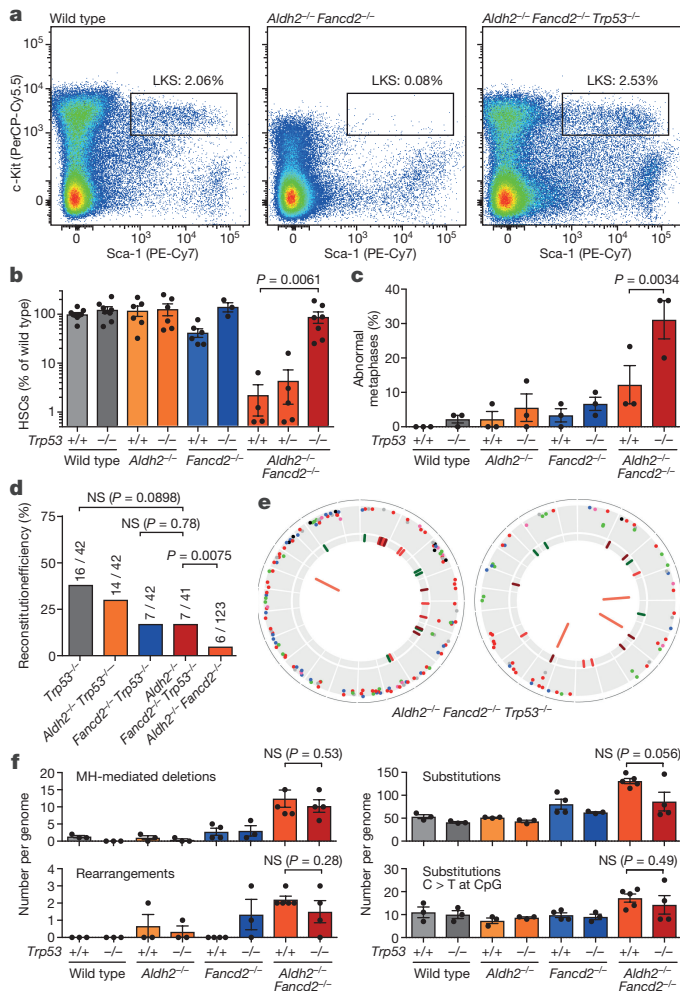


Figure 6 | A p53 response depletes aldehyde-damaged HSCs.

a, Representative flow cytometry plot of HSPCs (LKS). **b**, Quantification of HSCs as determined by flow cytometry (P calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; $n = 9, 8, 6, 6, 6, 3, 4, 5$ and 7 mice, left to right). **c**, Frequency of abnormal metaphases in bone marrow cells (P calculated by two-sided Fisher's exact test; data shown as mean and s.e.m.; 3 mice per genotype, 30 metaphases per mouse). See Extended Data Fig. 8b–d for a complete list of rearrangements. **d**, Proportion and number of irradiated recipients that were positive for reconstitution by transplantation of single HSCs (P calculated by two-sided Fisher's exact test). **e**, Mutations in two *Aldh2*^{−/−} *Fancd2*^{−/−} *Trp53*^{−/−} HSCs. **f**, Number of microhomology-mediated deletions, rearrangements, substitutions and clock substitutions per genome (P calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; $n = 3, 3, 3, 3, 4, 3, 5$ and 4 HSC genomes, left to right).

Discussion

These results outline the mechanisms by which the mouse haematopoietic system and, more specifically, blood stem cells respond to an endogenous and alcohol-derived source of DNA damage. Primary protection against acetaldehyde is provided by ALDH2-mediated detoxification and, when this is lost or saturated, acetaldehyde damages DNA. The Fanconi anaemia pathway is the principal mechanism to counteract this damage, but NHEJ and homologous recombination can also deal with these lesions. These results therefore illustrate that coordinated pathway choice is critical for maintaining genome stability upon aldehyde exposure. The Fanconi anaemia pathway prevents aldehyde lesions from degenerating into DSBs, the illegitimate repair of which leads to a characteristic pattern of mutagenesis in HSCs (Extended Data Fig. 11). Aldehydes are capable of forming a diverse range of DNA lesions—from base adducts to DNA–DNA or DNA–protein crosslinks. The known molecular function of the Fanconi

anaemia pathway suggests that the most physiologically toxic lesion caused by aldehydes may be a DNA interstrand crosslink. However, if it is indeed an interstrand crosslink, then the factors involved in the translesion synthesis or homologous recombination processes are distinct from the previously described mechanism of interstrand-crosslink repair^{6,14,15,28}. It will be important to resolve the nature of the lesion and the precise mechanics of its repair.

HSCs mutated by aldehydes are functionally compromised and display myeloid bias. The p53 response is critical in driving the loss in number and function of HSCs. Although *Trp53* deletion rescues HSC defects, this, paradoxically, does not result in further genomic instability at the single HSC level. It is important to emphasize, however, that the pool of HSCs is larger, and therefore there is still an overall increase in mutation. Nevertheless, our work implies that the relationship between p53, DNA repair and genome stability is more complex in stem cells than previously appreciated. The central role for ALDH2 in removing genotoxic aldehydes has implications for the more than 540 million people who are deficient in ALDH2 activity. Alcohol exposure in such individuals may cause DNA DSBs and chromosome rearrangements³. This large population may also be susceptible to alcohol-induced age-related blood disorders. More generally, this research provides a simple plausible explanation for the established epidemiological link between alcohol consumption and enhanced cancer risk^{4,29}.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 31 January; accepted 21 November 2017.

Published online 3 January 2018.

- Roswall, N. & Weiderpass, E. Alcohol as a risk factor for cancer: existing evidence in a global perspective. *J. Prev. Med. Public Health* **48**, 1–9 (2015).
- Wang, M. *et al.* Identification of DNA adducts of acetaldehyde. *Chem. Res. Toxicol.* **13**, 1149–1157 (2000).
- Lai, C. L. *et al.* Dominance of the inactive Asian variant over activity and protein contents of mitochondrial aldehyde dehydrogenase 2 in human liver. *Alcohol. Clin. Exp. Res.* **38**, 44–50 (2014).
- Yokoyama, A. & Omori, T. Genetic polymorphisms of alcohol and aldehyde dehydrogenases and risk for esophageal and head and neck cancers. *Jpn. J. Clin. Oncol.* **33**, 111–121 (2003).
- Garaycoechea, J. I. *et al.* Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature* **489**, 571–575 (2012).
- Langevin, F., Crossan, G. P., Rosado, I. V., Arends, M. J. & Patel, K. J. *Fancd2* counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* **475**, 53–58 (2011).
- Garaycoechea, J. I. & Patel, K. J. Why does the bone marrow fail in Fanconi anemia? *Blood* **123**, 26–34 (2014).
- Hira, A. *et al.* Variant ALDH2 is associated with accelerated progression of bone marrow failure in Japanese Fanconi anemia patients. *Blood* **122**, 3206–3209 (2013).
- Nijnik, A. *et al.* DNA repair is limiting for haematopoietic stem cells during ageing. *Nature* **447**, 686–690 (2007).
- Rossi, D. J. *et al.* Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature* **447**, 725–729 (2007).
- Milyavsky, M. *et al.* A distinctive DNA damage response in human haematopoietic stem cells reveals an apoptosis-independent role for p53 in self-renewal. *Cell Stem Cell* **7**, 186–197 (2010).
- Mohrin, M. *et al.* Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. *Cell Stem Cell* **7**, 174–185 (2010).
- Flach, J. *et al.* Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* **512**, 198–202 (2014).
- Niedzwiedz, W. *et al.* The Fanconi anaemia gene *FANCC* promotes homologous recombination and error-prone DNA repair. *Mol. Cell* **15**, 607–620 (2004).
- Long, D. T., Räschele, M., Joukov, V. & Walter, J. C. Mechanism of RAD51-dependent DNA interstrand cross-link repair. *Science* **333**, 84–87 (2011).
- Tan, S. L. W. *et al.* A class of environmental and endogenous toxins induces *BRCA2* haploinsufficiency and genome instability. *Cell* **169**, 1105–1118.e15 (2017).
- Bunting, S. F. *et al.* 53BP1 inhibits homologous recombination in *Brca1*-deficient cells by blocking resection of DNA breaks. *Cell* **141**, 243–254 (2010).
- Adamo, A. *et al.* Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. *Mol. Cell* **39**, 25–35 (2010).
- Pace, P. *et al.* Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. *Science* **329**, 219–223 (2010).
- Osawa, M., Hanada, K., Hamada, H. & Nakauchi, H. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science* **273**, 242–245 (1996).

21. Dykstra, B., Olthof, S., Schreuder, J., Ritsema, M. & de Haan, G. Clonal analysis reveals multiple functional defects of aged murine hematopoietic stem cells. *J. Exp. Med.* **208**, 2691–2703 (2011).
 22. Pang, W. W. *et al.* Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc. Natl Acad. Sci. USA* **108**, 20012–20017 (2011).
 23. Ghezraoui, H. *et al.* Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. *Mol. Cell* **55**, 829–842 (2014).
 24. Wyatt, D. W. *et al.* Essential roles for Polymerase θ -mediated end joining in the repair of chromosome breaks. *Mol. Cell* **63**, 662–673 (2016).
 25. García-Rubio, M. L. *et al.* The Fanconi anemia pathway protects genome integrity from R-loops. *PLoS Genet.* **11**, e1005674 (2015).
 26. Schwab, R. A. *et al.* The Fanconi anemia pathway maintains genome stability by coordinating replication and transcription. *Mol. Cell* **60**, 351–361 (2015).
 27. Jaber, S., Toufekhtchan, E., Lejour, V., Bardot, B. & Toledo, F. p53 downregulates the Fanconi anaemia DNA repair pathway. *Nat. Commun.* **7**, 11091 (2016).
 28. Oberbeck, N. *et al.* Maternal aldehyde elimination during pregnancy preserves the fetal genome. *Mol. Cell* **55**, 807–817 (2014).
 29. Kawashima, N. *et al.* Aldehyde dehydrogenase-2 polymorphism contributes to the progression of bone marrow failure in children with idiopathic aplastic anaemia. *Br. J. Haematol.* **168**, 460–463 (2015).
- V. Romashova and M. Balmont for help with flow cytometry; and J. Sale, C. Rada, M. Taylor, Y. L. Wu and members of the Patel laboratory for critical reading of the manuscript. The Human Research Tissue Bank (supported by the NIHR Cambridge Biomedical Research Centre) processed histology. K.J.P. is supported by the MRC and the Jeffrey Cheah Foundation. G.P.C. and L.M. were supported by CRUK. J.I.G. is supported by the Wellcome Trust and King's College, Cambridge.

Author Contributions J.I.G., G.P.C. and K.J.P. conceived the study and wrote the manuscript. J.I.G. conducted the majority of the experiments and analysed the data. G.P.C. assisted in the characterization of genomic instability in *Aldh2^{-/-}Fancd2^{-/-}* mice and single HSC transplantation. F.L. analysed the survival of chicken DT40 cells, performed western blotting and assisted with the analysis of micronucleus samples. L.M. assisted with single cell transplantation and performed the BigBlue *in vivo* point-mutation analysis. S.L. and F.Y. performed the M-FISH karyotyping of mouse metaphases. N.P. performed validations of indels by targeted deep sequencing. G.G., S.R., S.N.-Z. and M.R.S. provided assistance with the analysis and interpretation of sequencing data.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to K.J.P. (kjp@mrc-lmb.cam.ac.uk).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Kent for technical advice with single-HSC transplants; R. Berks, A. Middleton, J. Wiles, C. Knox, X. Gong, J. Roe, J. Willems, the ARES staff and Biomed for their help with mouse work; M. Daly, F. Zhang,

METHODS

Mice. *Aldh2*^{-/-}*Fancd2*^{-/-} mice were generated on a C57BL/6 × 129S4S6/Sv F₁ background. To this end, the previously reported *Fancd2* allele (*Fancd2*^{tm1Hou}; MGI ID: 2673422, a gift from M. Grompe) was back-crossed onto the C57BL/6Jol background for 10 generations and crossed with *Aldh2*^{+/-} (C57BL/6N) mice to generate *Aldh2*^{+/-}*Fancd2*^{+/-} mice on a pure C57BL/6 background. Likewise, the previously reported *Aldh2* allele (*Aldh2*^{tm1a(EUCOMM)Wtsi}; MGI ID: 4431566, EUCOMM⁶) was backcrossed from C57BL/6N onto 129S6/Sv for five generations and crossed with *Fancd2*^{+/-} mice to generate *Aldh2*^{+/-}*Fancd2*^{+/-} mice on a 129S4S6/Sv background. Finally, *Aldh2*^{-/-}*Fancd2*^{-/-} and control mice were generated as F₁ hybrids from crosses between *Aldh2*^{+/-}*Fancd2*^{+/-} females (129S4S6/Sv) and *Aldh2*^{+/-}*Fancd2*^{+/-} males (C57BL/6).

To generate *Fanca*^{-/-}*Ku70*^{-/-} mice on a pure C57BL/6 background, *Fanca*^{+/-} mice (*Fanca*^{tm1a(EUCOMM)Wtsi}; MGI ID: 4434431, C57BL/6N, EUCOMM⁵) were crossed with *Ku70*^{+/-} mice (*Xrcc6*^{tm1Fwa}; MGI ID: 2179954³⁰), and the *Fanca*^{+/-}*Ku70*^{+/-} progeny were then intercrossed to generate all possible genotypes. Pups from these crosses were genotyped at between two and three weeks old. For the generation of *Fanca*^{fl/fl}*Ku70*^{-/-}*Vav1-iCre*⁺ tissue-specific double mutants (also on a pure C57BL/6 background), *Fanca*^{+/-} mice were first crossed with FLP deleter mice³¹ to produce the *Fanca* floxed allele (*Fanca*^{fl} or *Fanca*^{tm1c(EUCOMM)Wtsi}). Recombination of the Frt sites was verified by PCR (using the primers FL033, GCCTTGTGCTGCTAATTCATGT; FL040, TCAGCTCACTGAGACGCAACCTTTT ACAT; and En2A, GCTTCACTGAGTCTCTGGCATCTC), and reconstitution of FANCA expression was verified by western blotting spleen extracts of *Fanca*^{fl/fl} mice (Extended Data Fig. 3). *Fanca*^{+/fl} mice were then crossed with *Ku70*^{+/-} mice to eventually produce *Fanca*^{fl/fl}*Ku70*^{+/-} mice. Finally, these mice were crossed with *Fanca*^{+/-}*Ku70*^{+/-} *Vav1-iCre* to generate *Fanca*^{fl/fl}*Ku70*^{-/-}*Vav1-iCre* and control mice. The *Vav1-iCre* allele directs the expression of the iCre recombinase to HSCs and haematopoietic tissues³², and in this case yields the *Fanca*-null allele (*Fanca*^Δ or *Fanca*^{tm1d(EUCOMM)Wtsi}). *Fanca*^{Δ/Δ} mice phenocopy the *Fanca*^{-/-} mice reported previously, as judged by FANCA expression, sterility and sensitivity to mitomycin C (Fig. 3, Extended Data Fig. 3).

Similarly to *Aldh2*^{-/-}*Fancd2*^{-/-} F₁ mice, *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} mice were also generated in a C57BL/6 × 129S4S6/Sv F₁ background. In brief, the *Trp53* allele reported previously³³ was backcrossed onto 129S6/Sv or C57BL/6J for six generations. *Trp53*^{+/-} mice were then intercrossed with *Aldh2*^{-/-}*Fancd2*^{+/-} mice to establish parental (F₀) strains on both genetic backgrounds, which were finally crossed to obtain *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} and control F₁ mice.

For single HSC transplantation experiments, we used CD45.1 homozygous mice on a C57BL/6J × 129S6/Sv F₁ background as recipients. CD45.1 (or *Ptprc*^u) had been serially backcrossed from B6.SJL onto 129S6/Sv for six generations, with selection at each generation by serotyping with anti-CD45.1 (A20, FITC, BioLegend) and anti-CD45.2 antibodies (104, PE-Cy7, BioLegend).

For the *in vivo* point-mutation assay, mice carrying BigBlue λ LIZ shuttle vector repeats (Stratagene) were crossed with *Aldh2*^{-/-}*Fancd2*^{+/-} mice on a C57BL/6 × 129S4S6/Sv hybrid background. The resulting mice were intercrossed to obtain *Aldh2*^{-/-}*Fancd2*^{-/-} BigBlue λ LIZ and control mice.

For the analysis of Mendelian segregation of alleles, sample size was determined by power analysis using <http://biomath.info/power/chsq1gp.htm>. Sufficient mice to detect a 50% reduction in expected frequency were used, using power of 0.8 and alpha 0.05. No statistical methods were used to predetermine sample size in the other animal experiments. No randomization was employed. The investigators were blinded to the genotypes of mice throughout the study and data were acquired by relying purely on identification numbers.

All animals were maintained in specific pathogen-free conditions. In individual experiments all mice were matched for gender and age (8–12 weeks). All animal experiments undertaken in this study were done so with the approval of the UK Home Office.

Ethanol treatment. For acute ethanol exposure, *Aldh2*^{-/-}*Fancd2*^{-/-} mice and appropriate controls were injected intraperitoneally with ethanol. The total dose of 5.8 g kg⁻¹ was split into two injections separated by 4 h. Ethanol (96%, Sigma) was diluted to 28% v/v in saline, and administered twice at 13 ml kg⁻¹. Mice were exsanguinated 48 h after the second injection and peripheral blood was analysed with the micronucleus assay. Alternatively, mice were injected with colchicine for the preparation of metaphase spreads for M-FISH. For SCE analysis, mice were injected with colchicine 12 h after the second ethanol injection.

For chronic ethanol treatment of *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} and control mice, ethanol was administered in drinking water for ten days as reported previously⁶. For the first five days, the drinking water supply was replaced by a solution of 10:15:75 blackcurrant Ribena:ethanol:water, followed by a 10:20:80 solution for the last five days. A 50 μ l blood sample was taken from tail veins before alcohol exposure, and by cardiac puncture at the end of the experiment, to measure full

blood counts. Femurs were dissected for histological analysis and to determine bone marrow cellularity.

Preparation of mouse bone marrow for metaphase spreads. Treated or untreated young mice (8–12 weeks of age) were injected intraperitoneally with 100 μ l of colchicine (0.5% w/v in water, Sigma). After 30 min the mice were culled by cervical dislocation, femurs were harvested and placed in ice-cold PBS. Bone marrow cells were flushed with 10 ml of pre-warmed hypotonic solution (75 mM KCl, 37°C) through a 70- μ m cell strainer and incubated for 15 min in a water bath at 37°C. After the incubation, 1 ml of fixative (3:1 methanol:acetic acid) was added dropwise to the hypotonic buffer, and mixed by gentle inversion of the tube. The tubes were spun down for 10 min at 250g and the supernatant was aspirated, leaving 50 μ l and the cell pellet in the tube. The cells were resuspended by flicking the base of the tube very gently, 3 ml of fixative were added dropwise and the volume was made up to 10 ml by pipetting fixative down the side of the tube. The cells were incubated at room temperature for 30 min and stored at -20°C until further use.

SCE assay for mouse bone marrow. The staining of metaphase spreads for the quantification of SCEs was adapted from published protocols^{34,35}. A 50 mg BrdU slow-release pellet (Innovative Research of America) was surgically implanted subcutaneously into 8-to-12-week-old mice. Unchallenged mice were injected with colchicine 24 h later and metaphases were prepared after 30 min. Mice challenged with ethanol were injected intraperitoneally with ethanol 8 h and 12 h after implantation of the BrdU pellet. A total ethanol dose of 5.8 g kg⁻¹ was split between these two doses as described previously. Metaphases were prepared as outlined above. Cells were then dropped from a height of 30 cm onto chilled, humidified slides. The slides were then dried for 1 h at 62°C in a hybridization oven. Cells were washed in 2 × SSC for 5 min at room temperature. Cells were stained for 15 min at room temperature with 1 μ g ml⁻¹ Hoechst 33258 pentahydrate (H3569, Molecular Probes) in 2 × SSC. The slides were then transferred to a Petri dish with 2 × SSC and exposed to UV irradiation for 30 min in a Stratagene Crosslinker (Stratagene). The slides were then dehydrated by passing them through an ethanol series (70%, 96% and 100%) and placed in PBS for 5 min at room temperature. The DNA was denatured by exposure to 0.07 N NaOH for 2 min at room temperature. The slides were then washed three times in PBS for 5 min. The slides were then blocked in PBS, 1% BSA, 0.5% Tween-20 for 1 h at room temperature and stained overnight with a FITC-conjugated mouse anti-BrdU antibody (Clone B44, BD Biosciences) diluted 1:1 in PBS, 3% BSA, 0.5% Tween-20 at room temperature. The slides were then washed three times with PBS, 1% BSA, 0.5% Tween-20 for 5 min at room temperature and stained with goat anti-mouse Alexa Fluor-488 secondary antibody (A-11001, Life Technologies) diluted 1:500 in PBS, 1% BSA, 0.5% Tween-20 for 6 h at room temperature. The slides were then washed three times in PBS, 1% BSA, 0.5% Tween-20 for 15 min and stained with Hoechst 33342 trihydrochloride (H3570, Molecular Probes) diluted 1:2000 in PBS for 15 min at room temperature. The slides were then washed three times in PBS for 10 min on each occasion, washed once in water for 5 min, mounted with ProLong Gold Antifade Mountant (P36930, Molecular Probes) and coverslips were lowered onto the slides. Thirty metaphases were captured per sample using a Zeiss LSM 780 confocal microscope (Zeiss). The number of sister-chromatid exchanges per metaphase was then counted blind.

M-FISH karyotyping. For M-FISH, chromosome-specific DNA libraries were generated from flow-sorted chromosomes provided by the Flow Cytometry Core Facility of the Wellcome Trust Sanger Institute, using the GenomePlex Complete whole-genome amplification kit (Sigma-Aldrich). A mouse 21-colour painting probe was prepared following the pooling strategy³⁶. Five mouse-chromosome pools were each labelled with ATTO 425-, ATTO 488-, Cy3-, Cy5- and Texas Red-dUTPs (Jena Bioscience), respectively, using the GenomePlex WGA reamplification kit (Sigma-Aldrich) and a dNTP mixture as described previously³⁷. The labelled products were pooled and sonicated to achieve a size range of 200–1,000 bp, optimal for use in chromosome painting. The sonicated DNA was ethanol-precipitated together with mouse Cot-1 DNA (Thermo Fisher Scientific), and resuspended in a hybridization buffer (50% formamide, 2 × SSC, 10% dextran sulfate, 0.5 M phosphate buffer, 1 × Denhardt's solution, pH 7.4). Bone marrow cells suspended in fixative as described above (3:1 methanol:acetic acid) were dropped onto pre-cleaned microscope slides, followed by fixation in acetone (Sigma-Aldrich) for 10 min and dehydration through an ethanol series (70%, 90% and 100%). Metaphase spreads on slides were denatured by immersion in an alkaline denaturation solution (0.5 M NaOH, 1.0 M NaCl) for 2 min, followed by rinsing in 1 M Tris-HCl (pH 7.4) solution for 3 min, 1 × PBS for 3 min and dehydration through a 70%, 90% and 100% ethanol series. The M-FISH probes were denatured at 65°C for 10 min before being applied onto the denatured slides. The hybridization area was sealed with a 22 mm × 22 mm coverslip and rubber cement. Hybridization was carried out in a 37°C incubator for approximately 44–48 h. The post-hybridization washes included a 5-min stringent wash in 0.5 × SSC at 75°C, followed by a 5 min rinse in 2 × SSC containing 0.05% Tween-20 (VWR) and a

2 min rinse in $1 \times$ PBS, both at room temperature. Finally, slides were mounted with SlowFade Gold mounting solution containing DAPI (Thermo Fisher Scientific). M-FISH images were visualized on a Zeiss AxioImager D1 fluorescent microscope equipped with narrow band-pass filters for DAPI, DEAC, FITC, Cy3, Texas Red and Cy5 fluorescence and an ORCA-EA CCD camera (Hamamatsu). M-FISH digital images were captured using the SmartCapture software (Digital Scientific UK), and processed using the SmartType Karyotyper software (Digital Scientific UK). Thirty metaphases for each sample were karyotyped by M-FISH.

DT40 clonogenic survival. DT40 cells were grown in RPMI Medium 1640 (Life Technologies, 61870), supplemented with 7% fetal bovine serum (FBS, Life Technologies, 10270), 3% chicken serum (Life Technologies, 16110), $50 \mu\text{M}$ β -mercaptoethanol and penicillin/streptomycin, at 37°C in a 5% CO_2 incubator. Sensitivity assays were performed as previously described⁶. In brief, 10^5 cells were incubated with drug-containing medium in a sealed FACS tube at 37°C for 2 h (acetaldehyde) or 1 h (mitomycin C or cisplatin). Dilutions were plated in 6-well plates containing semi-solid medium (4000 cP methyl cellulose (M0512 Sigma), DMEM/F-12 powder (Life Technologies, 32500-043), 7% FBS, 3% chicken serum, $50 \mu\text{M}$ β -mercaptoethanol and penicillin/streptomycin). Plates were incubated for 7–10 days, after which time colonies were counted manually. Survival is plotted as a percentage relative to untreated cells. Each data point represents the mean of three independent experiments each carried out in quadruplicate.

Sensitivity assays of primary mouse B cells. These assays were performed as described previously⁵. Lymphocytes purified from the spleen using Lympholyte M (Cederlane) were stimulated with lipopolysaccharide (L4391, Sigma) at a final concentration of $40 \mu\text{g ml}^{-1}$. Cells (4×10^5) were then plated with acetaldehyde in one well of a 24-well plate. After seven days, viable cells were counted using trypan blue exclusion from 100 images on a Vi-Cell XR cell viability counter (Beckman Coulter). Each data point represents the mean of three independent experiments each carried out in quadruplicate.

Survival assays of colony-forming units (CFU). Bone marrow cells were isolated using IMDM medium, and single cell suspensions were obtained by passing the bone marrow through a $70\text{-}\mu\text{m}$ cell strainer (Falcon). Nucleated cells were counted by diluting cells tenfold in a 3% solution of acetic acid with methylene blue (Stem Cell Technologies) using a Vi-Cell XR cell viability counter (Beckman Coulter). Cells were resuspended to make up 1.5 ml of IMDM containing 30×10^6 cells and $250 \mu\text{l}$ of each suspension was mixed with $250 \mu\text{l}$ of IMDM containing $2 \times$ acetaldehyde to give final concentrations of 0, 1, 2, 4 and 8 mM acetaldehyde. The cells were incubated at 37°C for 4 h in sealed tubes, after which two tenfold serial dilutions were made. $400 \mu\text{l}$ of cells were then added to 4 ml of MethoCult M3534 (StemCell Technologies), and the total volume of each dilution was plated in two wells of a six-well plate each containing 10^6 , 10^5 and 10^4 cells, respectively. After seven days of culture at 37°C with 5% CO_2 , the colonies were counted and the relative survival was plotted. Each data point represents the average of experimental duplicates carried out on three mice of each genotype.

Flow cytometry. The micronucleus assay was performed essentially as described previously³⁸. Treated or untreated mice (8–12 weeks of age) were bled and $6 \mu\text{l}$ blood was mixed with $338 \mu\text{l}$ PBS supplemented with $1,000 \text{ U ml}^{-1}$ of heparin (Calbiochem). $360 \mu\text{l}$ of blood suspension was then added to 3.6 ml of methanol at -80°C and stored at -80°C for at least 12 h. 1 ml of fixed blood cells was then washed with 6 ml of bicarbonate buffer (0.9% NaCl, 5.3 mM NaHCO_3). The cells were resuspended in $150 \mu\text{l}$ of bicarbonate buffer and $20 \mu\text{l}$ of this suspension was used for subsequent staining. $72 \mu\text{l}$ of bicarbonate buffer, $1 \mu\text{l}$ of FITC-conjugated CD71 antibody (GenTex, clone R17217.1.4) and $7 \mu\text{l}$ RNase A (Sigma) were premixed and added to $20 \mu\text{l}$ of each cell suspension. The cells were stained at 4°C for 45 min, followed by addition of 1 ml bicarbonate buffer and centrifugation. Finally, cell pellets were resuspended in $500 \mu\text{l}$ bicarbonate buffer supplemented with $5 \mu\text{g ml}^{-1}$ propidium iodide (Sigma). The samples were analysed immediately on an LSRII FACS analyser (BD) and the data analysed with FlowJo v10.0.7.

For HSC quantification, bone marrow cells were isolated from tibiae and femurs with staining buffer (PBS supplemented with 2.5% FCS) and strained through $70\text{-}\mu\text{m}$ meshes. Red cells were lysed by resuspending the cells in 10 ml red cell lysis buffer (MACS Miltenyi Biotec) for 10 min at room temperature. After centrifugation, the cell pellet was resuspended in staining buffer and nucleated cells were counted with 3% acetic acid (StemCell Technologies) on a Vi-Cell XR cell viability counter (Beckman Coulter). Bone marrow cells (10×10^6 cells) were resuspended in $200 \mu\text{l}$ of staining buffer containing the following antibody solution: FITC-conjugated lineage cocktail with antibodies against CD4 (clone H129.19, BD Pharmingen), CD3e (clone 145-2C11, eBioscience), Ly-6G/Gr-1 (clone RB6-8C5, eBioscience), CD11b/Mac-1 (clone M1/70, BD Pharmingen), CD45R/B220 (clone RA3-6B2, BD Pharmingen), Fc ϵ R1 α (clone MAR-1, eBioscience), CD8a (clone 53-6.7, BD Pharmingen), CD11c (clone N418, eBioscience), TER-119 (clone Ter119, BD Pharmingen) and CD41 (FITC, clone MWReg30, BD Pharmingen); c-Kit (PerCP-Cy5.5, clone 2B8, eBioscience), Sca-1 (PE-Cy7, clone

D7, eBioscience), CD150 (PE, clone TC15-12F12.2, BioLegend) and CD48 (biotin, clone HM48-1, BioLegend). The samples were incubated for 15 min at 4°C and washed with 2 ml buffer. The cell pellets were resuspended in $200 \mu\text{l}$ staining buffer containing streptavidin-BV421 and incubated for another 15 min at 4°C . Finally, cells were washed, resuspended in $500 \mu\text{l}$ staining buffer, data were acquired on a Fortessa FACS analyser (Becton Dickinson) and analysed with FlowJo v10.0.7. LKS cells were defined as lineage $^-$ CD41 $^-$ Sca-1 $^+$ Kit $^+$ and HSCs were defined as LKS CD48 $^-$ CD150 $^+$.

To assess engraftment of single HSCs into irradiated recipients, $50 \mu\text{l}$ of blood was obtained from the tail vein of recipient mice every two weeks. RBCs were lysed by the addition of 1 ml of ammonium chloride lysis buffer ($155 \text{ mM NH}_4\text{Cl}$, 10 mM KHCO_3 , $0.1 \text{ mM Na}_2\text{EDTA}$, pH 7.2) and incubated for 10 min at room temperature. After centrifugation, the cell pellets were resuspended in $100 \mu\text{l}$ of staining buffer containing antibodies against: CD4 (FITC, clone H129.19, BD Pharmingen), CD8a (FITC, clone 53-6.7, BD Pharmingen), CD45R/B220 (PerCP-Cy5.5, clone RA3-6B2, BioLegend), CD11b/Mac-1 (PE, clone M1/70, BD Pharmingen), Ly-6G/Gr-1 (PE, clone 1A8, BD Pharmingen), TER-119 (PE-Cy7, clone TER-119, BioLegend), CD45.1 (BV421, clone A20, BioLegend) and CD45.2 (APC, clone 104, BioLegend). After incubation, cells were washed with 3 ml of staining buffer before being resuspended in $250 \mu\text{l}$ of the same buffer. Samples were run on the Fortessa analyser (BD) with the HTS module and the multilineage chimaerism was calculated using FlowJo v10.0.7. TER-119 was used to exclude RBC debris and chimaerism was calculated for each of the WBC lineages (CD45 $^+$ total WBCs, B220 $^+$ B cells, CD4 $^+$ CD8 $^+$ T cells and Gr-1 $^+$ Mac-1 $^+$ myeloid cells) as the proportion of cells derived from the single HSC (CD45.2 $^+$) over the total number of CD45 $^+$ cells, which includes cells derived from the recipient or carrier cells (CD45.1 $^+$).

For intracellular staining of p53 and cleaved caspase-3, total bone marrow cells were stained with the lineage-depletion kit (130-090-858, MACS Miltenyi Biotec) following the manufacturer's instructions and passed through LS magnetic columns. Lineage-depleted cells were spun down for 5 min at 1,200 r.p.m. and the pellets were resuspended in $200 \mu\text{l}$ MACS buffer with the antibodies described above for HSC quantification. In parallel, 3×10^6 total bone marrow cells were stained with antibodies against committed lineages: CD45R/B220 (PE, clone RA3-6B2, BD Pharmingen) and IgM (FITC, clone II/41, BD Pharmingen) for B cell progenitors; TER-119 (FITC, clone TER-119, BD Pharmingen) and CD71 (PE, clone C2, BD Pharmingen) for erythroid maturation; and CD11b/Mac-1 (PE, clone M1/70, BD Pharmingen) and Ly-6G/Gr-1 (FITC, clone 1A8, eBioscience) for monocyte/granulocyte progenitors. After antibody staining, cells were washed, then fixed and permeabilized with BD Cytofix/Cytoperm solution (554722, BD Pharmingen) following the manufacturer's instructions. Finally, cells were stained with either anti-p53 (AlexaFluor647, clone 1C12, Cell Signalling) or anti-cleaved-caspase-3 (AlexaFluor647, clone D3E9, Cell Signalling) antibodies.

Gating strategies are described in Supplementary Fig. 2.

Blood counts. Total blood was collected in K3EDTA MiniCollect tubes (Greiner bio-one) and analysed on a VetABC analyser (Horiba).

Western blot. The FANCA antibody (Cell Signalling, D1L2Z) was used at 1:1000 in 5% w/v BSA, $1 \times$ TBS, 0.1% Tween-20 at 4°C with gentle shaking, overnight. The β -actin antibody (Abcam, 8227) was used at 1:3,000 in the same conditions. Swine anti-rabbit immunoglobulins HRP (Dako) was used as secondary antibody at 1:2,000 for 1 h at room temperature.

Histological analysis. Tissue samples were fixed in 10% neutral-buffered formalin for at least 24 h. The femur samples were then decalcified and embedded in paraffin, and $4 \mu\text{m}$ sections were cut before staining with haematoxylin and eosin using standard methods.

In vivo point mutation assay. The λ select-cII (BigBlue) mutagenesis assay was performed following the manufacturer's instructions. This assay allows the detection of point mutations *in vivo* and is based on the ability of coliphage λ to multiply either through the lytic or lysogenic cycles in *Escherichia coli*.

In brief, the RecoverEase DNA-isolation kit (Stratagene) was used to extract genomic DNA from the bone marrow of 8-to-12-week old *Aldh2* $^{-/-}$ *Fancd2* $^{-/-}$ and control mice that carry Big Blue λ LIZ repeats. The shuttle vector that was recovered from mouse genomic DNA was packaged into phage with the Transpack packaging extract (Stratagene), which was then used to infect *E. coli* G1250 (Stratagene).

To assess the frequency of mutations within the *cII* gene, infected *E. coli* were diluted in TB1 top agar, spread on ten 100-mm TB1-agar plates and incubated at 24°C for 48 h. The plaques were enumerated, picked and replated to confirm that they were lytic. This provided the number of mutated phage within the sample. To determine the total number of phage undergoing the lysogenic cycle and the mutation frequency, 10- and 50-fold dilutions of the stock of infected *E. coli* in TB1 top agar were spread on two 100-mm TB1-agar plates and incubated at 37°C for 24 h. Incubation at 37°C switches all phage to the lytic cycle, and therefore allows the total number of replication-competent phage to be assessed. The mutation

frequency was then calculated as the number of *cII*-mutated plaques (24°C)/total number of plaques (37°C).

As a positive control for the *in vivo* λ select-*cII* mutagenesis assay, mice were exposed to 150 mg kg⁻¹ N-ethyl-N-nitrosourea (ENU, Sigma) in a single intraperitoneal injection on one occasion. Mice were allowed to recover and were euthanized 3 weeks later.

CFU-S₁₂ assays. CFU-S₁₂ assays were performed as described previously⁵, except that CFU-S colonies were counted after 12 days. In brief, to assess the frequency of CFU-S in mutant mice, total bone marrow was flushed from the femora and tibiae of mutant mice and appropriate controls. Nucleated cells were enumerated using a solution of 3% acetic acid and methylene blue and injected intravenously into 20 recipient mice that had been lethally irradiated. After 12 days the spleens were fixed in Bouin's solution (Sigma), and the number of colonies were counted and expressed relative to the number of total bone marrow cells injected.

To assess the survival of CFU-S₁₂ after exposure to acetaldehyde, we treated total bone marrow cells with 4 mM acetaldehyde for 4 h *in vitro* before injecting them into lethally irradiated recipient mice. After 12 days, the number of CFU-S were counted. Survival was expressed relative to the untreated control for each genotype. Each data point represents the mean CFU-S survival in ten recipient mice, expressed relative to untreated samples of the corresponding genotype.

Single HSC transplants. The single-stem-cell transplants were performed as described previously^{39,40}. CD45.1 recipient mice (C57BL/6J \times 129S6/Sv F1, 8-to-12-weeks old) were fed with water supplemented with the antibiotic enrofloxacin (Baytril, Bayer Corporation) for seven days before irradiation and for the duration of the experiment. The lethal radiation was delivered as a split dose of 1,000 rad (500 rad each, 3 h apart) using a ¹³⁷Cs GSR C1m source (GSR GmbH).

The lethally irradiated recipients were injected with single HSCs sorted from *Aldh2*^{-/-} *Fancd2*^{-/-} and control mice on a C57BL/6J \times 129S4S6/Sv F1 background (CD45.2, 8-to-12 weeks old). Bone marrow cells from these mice were extracted with IMDM medium (GIBCO), filtered through a 70- μ m strainer, spun down for 5 min at 1,200 r.p.m. and resuspended in 6 ml IMDM medium at room temperature. These cells were then overlaid onto 6 ml Lympholyte M (Cederlane) in a 15-ml Falcon tube and spun down for 20 min at 1,400g at room temperature with the brake off. The interface containing the mononucleated cells was transferred into another 15-ml tube and topped up with ice-cold MACS buffer (PBS pH 7.2, 0.5% BSA, 2 mM EDTA). After 5 min of centrifugation at 1,200 r.p.m., the cell pellets were resuspended in 320 μ l of MACS buffer, stained with the lineage-depletion kit (130-090-858, MACS Miltenyi Biotec) following the manufacturer's instructions and passed through LS magnetic columns. Lineage-depleted cells were spun down for 5 min at 1,200 r.p.m. and the pellets were resuspended in 200 μ l MACS buffer with the antibodies described previously (see Flow cytometry), except that anti-CD48 was directly conjugated to BV421 (clone HM48-1, BioLegend). The cells were resuspended in 500 μ l of MACS buffer and run on a Synergy sorter (Sony Biotechnology Inc.).

Single HSCs, defined as lineage⁻c-Kit⁺Sca-1⁺CD41⁻CD48⁻CD150⁺, were sorted into 100 μ l StemSpan SFEM medium (StemCell Technologies) in each well of a round bottom, 96-well plate (Costar) using a Synergy cell sorter (Sony Biotechnology). The plates were spun down for 5 min at 180g and the presence of a single cell per well was confirmed visually. The full content of selected wells was loaded into insulin syringes (29G, 0.5-inch needle) containing 2 \times 10⁵ carrier cells in 300 μ l Hank's balanced salt solution (StemCell technologies). The contents of the syringe were used to dislodge the single HSC from the bottom of the well, avoiding the creation of bubbles. The entire volume (400 μ l) was injected into the tail veins of irradiated recipients and chimaerism was measured every 2 weeks using flow cytometry. Recipients were considered reconstituted by the single stem cell if chimaerism for CD45.2⁺ WBCs was \geq 0.1%.

Whole-genome sequencing of mouse HSC clones. Single HSCs were allowed to expand *in vivo* for four months to guarantee that the transplanted cell was a stem cell. Recipient mice that were positive for reconstitution were euthanized four months after transplantation, and the blood, bone marrow, spleen and thymus were collected. All tissues were prepared for flow cytometry as described above and stained with CD45.1 (FITC, clone A20, BioLegend) and CD45.2 (APC, clone 104, BioLegend) antibodies. CD45.2⁺ cells were sorted from each tissue on a Synergy cell sorter (Sony Biotechnology), spun down and frozen at -80°C. Genomic DNA was then extracted from the CD45.2⁺ bone-marrow cells and from a tail biopsy that had been collected at 2 weeks of age from the same mouse that provided the single HSC. Genomic DNA was extracted with the Puregene Cell and Tissue kit (Qiagen) following the manufacturer's instructions.

Whole-genome sequencing was performed as described previously⁴¹. In brief, short-insert 500-bp genomic libraries were constructed according to Illumina library protocols and 100-base paired-end sequencing was performed on HiSeq 2000 or HiSeq X genome analysers to an average of 20 \times coverage. Short-insert paired-end reads were aligned to the reference mouse genome (NCBIM38) using

BWA-MEM (<http://bio-bwa.sourceforge.net/>). For each HSC clone, the matched tail sample was used as the reference. We sequenced two of the *Aldh2*^{-/-} *Fancd2*^{-/-} HSC clones to 40 \times coverage, together with their matched germ-line references. We noted a big overlap between the variants found at 20 \times and 40 \times coverage (data not shown), showing that doubling the coverage did not actually uncover many more mutations. Any additional calls found when the coverage was increased to 40 \times were predominantly subclonal, with a mean VAF of 0.22. Therefore, we concluded that 20 \times whole-genome sequencing provided sufficient coverage to allow us to uncover mutations present in the transplanted HSC.

Substitutions, indels and structural changes were called with the CaVEMan, Pindel and BRASS algorithms, which are described in detail elsewhere⁴². In addition to previously reported filtering⁴¹, we asked that all variants were unique to each HSC clone and not found in unrelated HSC or tail samples, or in unrelated mouse strains. We did not apply a VAF filter for clonality because when we examined the VAF distribution of the final datasets, these were centered around 0.5 (Extended Data Fig. 9). For the transcriptional analysis, previously reported HSC RNA-seq data was aligned with Bowtie2 (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>) against NCBIM38⁴³, and the overlap between indels and transcribed genes was calculated in R. To calculate the fraction of the genome covered by genes, positions of genes were retrieved from Ensembl. Overlapping regions between two genes were taken into account and counted only once.

Validation of indel and rearrangement calls. Indel calls of less than 50 bp were validated using multiplex PCR and targeted re-sequencing. We used 13 multiplex-primer combinations to capture and simultaneously amplify 172 (50%) of the previously identified indels. Primers were designed using MPprimer⁴⁴ to capture regions of between 190 and 250 bp in size (sequences available upon request). The first of round multiplex-PCR amplifications was performed with tailed gene primers and was individually barcoded by a second round of PCR with pre-validated MiSeq-ready primers⁴⁵, using a high fidelity polymerase (Q5 Hot Start HF, New England Biolabs). The PCR reaction conditions were as follows: Group 1, 100 ng DNA input in 25 μ l PCR reaction, 95°C for 2 min, six cycles of 98°C for 20 s, 65°C for 60 s, 60°C for 60 s, 55°C for 60 s, 50°C for 60 s and 70°C for 60 s, the reaction was then held at 4°C until addition of barcoded second-round primers, followed by 19 cycles of 98°C for 20 s, 62°C for 15 s and 72°C for 30 s, then 72°C 60 s; Group 2, 10 ng DNA input in 5 μ l PCR reaction, 95°C for 2 min, seven cycles of 95°C for 20 s, 58°C for 17 min and 70°C for 60 s, then held at 4°C until addition of barcoded second-round primers, followed by 23 cycles of 98°C for 20 s, 62°C for 15 s and 72°C for 30 s, then 72°C 60 s. Each sample was pooled, size selected by SPRI (\times 0.8) and quantified before being stored at -20°C until sequencing. Two MiSeq runs (300-bp paired-end) were used for variant confirmation, and reads were mapped with BWA. We analysed positions where the coverage was higher than 100 \times (159/172). With this approach, we were able to validate 91.2% of the original calls: 14/159 (8.8%) calls had VAF values <0.1 and were deemed false positives. The normal distribution of VAFs around 0.5 is consistent with most indels being clonal in origin.

For validation of rearrangements, we designed nested PCRs surrounding the breakpoints determined by the BRASS algorithm (primer sequences available upon request). PCR reactions were carried out in 20 μ l, using 10 ng (HSC clones or tails) or 400 ng (donor bone marrow) of genomic DNA and GoTaq G2 Hot Start Polymerase (M7401, Promega). The first round of PCR amplifications was performed at 95°C for 2 min, 35 cycles of 95°C for 30 s, 55°C for 20 s and 72°C for 30 s, then 72°C for 5 min. The reactions were diluted to 1 in 50 and 1 μ l of the diluted reaction was used as template for a second round of PCR with nested primers, to increase specificity and sensitivity, performed at 95°C for 2 min, 35 cycles of 95°C for 30 s, 60°C for 20 s and 72°C for 30 s, then 72°C for 5 min. The reactions were analysed on 2% agarose gels, bands of the expected sizes were excised and the identities of all products were confirmed by Sanger sequencing.

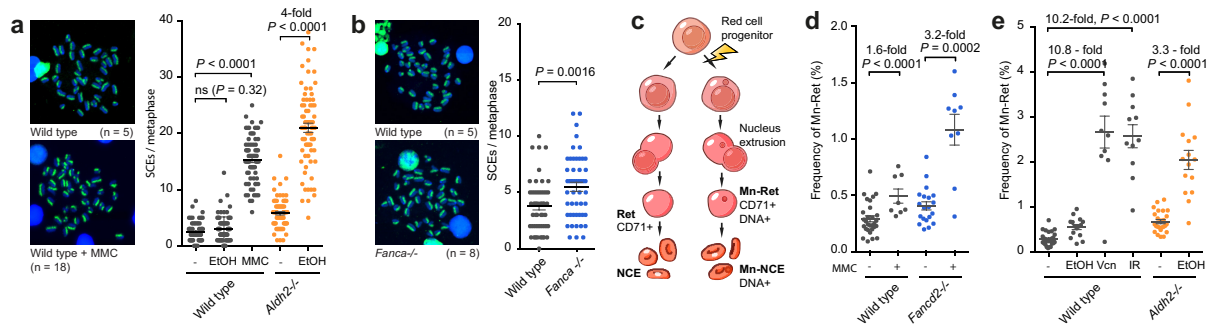
Using this approach, we found that 16/27 (59%) of rearrangements could be detected in the bone marrow of donor mice at the time the HSCs were transplanted (Extended Data Fig. 10). Any rearrangements present before transplantation must be clonal (that is, would not have arisen after the transplant). The failure to amplify the remaining 11 rearrangements by PCR does not mean that these are sub-clonal (that is, post-transplant) events. PCR amplification will depend on how much the transplanted HSC was contributing to blood production in the donor animal at the time of the transplant, as well as the sensitivity of each PCR. Therefore, we inferred clonality for the remaining calls by looking at loss of copy number (in the case of deletions, see Fig. 5l) and the number of reads involved in the rearrangement at the breakpoint for copy number-neutral changes.

Statistical analysis. Sample number (*n*) indicates the number of independent biological samples in each experiment. Sample numbers and experimental repeats are indicated in figure legends or Methods. Normality of data distribution was tested using the D'Agostino-Pearson omnibus normality test and variance was estimated before deciding on a statistical test. Unless otherwise stated in the

figure legend, data are shown as the mean \pm s.e.m. and the two-sided nonparametric Mann–Whitney test was used to assess statistical significance. Analysis was performed using GraphPad Prism.

Data availability. Whole-genome sequencing data have been deposited in the EMBL European Nucleotide Archive (ENA) under the accession code ERP009447. All other data are available upon reasonable request from the authors.

30. Gu, Y. *et al.* Growth retardation and leaky SCID phenotype of Ku70-deficient mice. *Immunity* **7**, 653–665 (1997).
31. Farley, F. W., Soriano, P., Steffen, L. S. & Dymecki, S. M. Widespread recombinase expression using FLP_{ER} (flipper) mice. *Genesis* **28**, 106–110 (2000).
32. de Boer, J. *et al.* Transgenic mice with hematopoietic and lymphoid specific expression of Cre. *Eur. J. Immunol.* **33**, 314–325 (2003).
33. Donehower, L. A. *et al.* Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356**, 215–221 (1992).
34. Giri, S. D. & Chatterjee, A. Modulation of mitomycin C-induced sister chromatid exchanges and cell cycle delay by buthionine sulfoximine and reduced glutathione in mouse bone marrow cells *in vivo*. *Mutat. Res.* **413**, 227–234 (1998).
35. Orsburn, B. *et al.* Differential requirement for H2AX and 53BP1 in organismal development and genome maintenance in the absence of poly(ADP)ribosyl polymerase 1. *Mol. Cell. Biol.* **30**, 2341–2352 (2010).
36. Geigl, J. B., Uhrig, S. & Speicher, M. R. Multiplex-fluorescence in situ hybridization for chromosome karyotyping. *Nat. Protoc.* **1**, 1172–1184 (2006).
37. Gribble, S. M. *et al.* Massively parallel sequencing reveals the complex structure of an irradiated human chromosome on a mouse background in the Tc1 model of Down syndrome. *PLoS ONE* **8**, e60482 (2013).
38. Reinholdt, L., Ashley, T., Schimenti, J. & Shima, N. Forward genetic screens for meiotic and mitotic recombination-defective mutants in mice. *Methods Mol. Biol.* **262**, 87–107 (2004).
39. Kent, D., Dykstra, B. & Eaves, C. Isolation and assessment of long-term reconstituting hematopoietic stem cells from adult mouse bone marrow. *Curr. Protoc. Stem Cell Biol.* (2007).
40. Oguro, H., Ding, L. & Morrison, S. J. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* **13**, 102–116 (2013).
41. Behjati, S. *et al.* Genome sequencing of normal cells reveals developmental lineages and mutational processes. *Nature* **513**, 422–425 (2014).
42. Nik-Zainal, S. *et al.* The genome as a record of environmental exposure. *Mutagenesis* **30**, 763–770 (2015).
43. Cabezas-Wallscheid, N. *et al.* Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell* **15**, 507–522 (2014).
44. Shen, Z. *et al.* MPprimer: a program for reliable multiplex PCR primer design. *BMC Bioinformatics* **11**, 143 (2010).
45. Quail, M. A. *et al.* A large genome center's improvements to the Illumina sequencing system. *Nat. Methods* **5**, 1005–1010 (2008).
46. Krishna, G., Fiedler, R. & Theiss, J. C. Simultaneous evaluation of clastogenicity, aneugenicity and toxicity in the mouse micronucleus assay using immunofluorescence. *Mutat. Res.* **282**, 159–167 (1992).



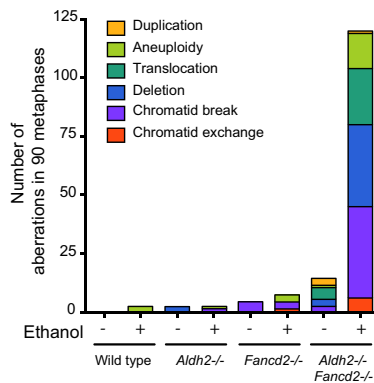
f Untreated

Wild type	0/30
	0/30
	0/30
<i>Aldh2</i> ^{-/-}	0/30
	0/30
	2/30 40,XX, Del(2)
	40,XX, Del(13)
<i>Fancc2</i> ^{-/-}	0/30
	1/30 40,XY, Chtb(X), Chtb(4)
	2/30 40,XY, Chtb(2)
	40,XY, Chtb(5)
<i>Aldh2</i> ^{-/-} <i>Fancc2</i> ^{-/-}	7/30 40,XY, Der(X)T(X;2), Der(2)T(X;2)
	40,XY, T(12;17), Dp(16), Del(16)
	40,XY, Der(8)T(8;14)
	40,XY, Dp(2)
	40,XY, Dp(2), Chtb(1)
	40,XY, Del(6)
	40,XY, Der(1)T(1;5), Der(5)T(1;5), Der(13)T(2;13)
	2/30 40,XY, Dp(16)
	40,XY, Chtb(X)
	2/30 40,XY, Del(2)
	39,XY, -19, Rb(12;19)

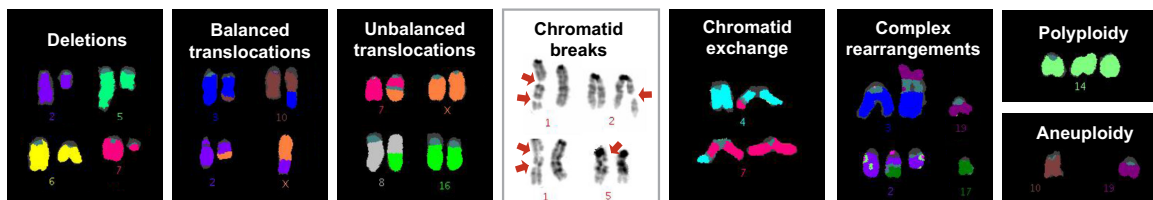
g 48 hours after Ethanol IP (5.8 g/kg)

Wild type	2/30 39,X,-Y
	39,X,-Y
	0/30
	0/30
<i>Aldh2</i> ^{-/-}	0/30
	0/30
	1/30 41,XX,+9, Chtb(9)
<i>Fancc2</i> ^{-/-}	4/30 40,XY, Chtb(4)
	40,XY, Chte(11;4)x1
	41,XY, +11
	41,XY, +7
<i>Aldh2</i> ^{-/-} <i>Fancc2</i> ^{-/-}	22/30 40,XY, Del(13)
	39,XY, -8, Del(2), T(8;15)
	40,XY, -1, Del(7), Del(17), +Del(15)
	40,XY, -2, +T(3;11)
	40,Y, Del(X), Del(6), T(7;11)
	40,XY, Del(1), Der(13)T(2;13)
	40,XY, Chte(5;18)x2, Chte(2;12)
	40,XY, Chte(9;17), Del(5), Chtb(8), Chtb(9)
	40,XY, Chte(13;14), Chtb(3)
	40,XY, -9, Chtb(15)
	40,XY, Del(12), Chtb(7), Chtb(12), Chtb(13)
	40,XY, Del(2), Der(17)T(16;17), Chtb(10)
	40,XY, Dp(4), Del(5), Chtb(2)
	40,XY, Del(X), Chtb(4)
	40,XY, Del(5), Chtb(14)
	41,XY, -6, +15, Del(14), Chtb(15), Der(X)T(6;X)
	40,XY, Chtb(1), Chtb(14)
	40,XY, Chtb(12)
	40,XY, Chtb(6)
	40,XY, Chtb(4)
	40,XY, Chtb(9)
	40,XY, Chtb(6)
	17/30 40,XX, Der(6)T(1;6)
	40,XX, Del(12)
	40,XX, Chte(4;7)x2
	40,XX, Del(6), Der(12)T(6;12)
	40,XX, Del(2), Chtb(13)
	40,XX, Del(13), Der(16)T(1;16)
	40,XX, Chtb(5)
	40,XX, Chtb(1)
	39,XX, -12, Del(4), Der(7)T(7;15), Der(15)T(7;15)
	40,XX, Del(2), Chtb(7)
	40,XX, Chtb(19)
	40,XX, Del(X)
	40,XX, Del(7), Chtb(1); Chtb(9); Chtb(16)
	39,XX, -11, Del(10), Der(14)T(10;14)
	40,XX, Del(4)
	40,XX, Chtb(1)
	40,XX, Del(1)
	39,XX, Der(8)T(8;18), Der(10;19)
	40,XX, Del(5), Der(8)T(8;16), Chtb(X)
	40,XX, Del(14), Der(18)T(6;18)
	40,XX, Del(6)
	40,XX, Chtb(2)
	40,XX, T(X;7), Del(2)
	40,XX, -6, +Del(17), Chtb(14)
	40,XX, -17, +Del(2), Der(2)T(2;17)
	40,XX, Del(7), +14, Chtb(16)
	40,XX, Der(7)T(4;7), +Del(8), ...
	...Chtb(1), Chtb(2), Chtb(15)
	40,XX, Del(X), Der(3)T(3;10), ...
	...Der(10)T(3;10), Chtb(2), Chtb(3), Chtb(10)
	39,XX, -1, Der(12)T(9;12)
	40,XX, Del(10), Chtb(6)
	40,XX, Der(3;19), -19
	40,XX, -10, dic(18;18)

h



i

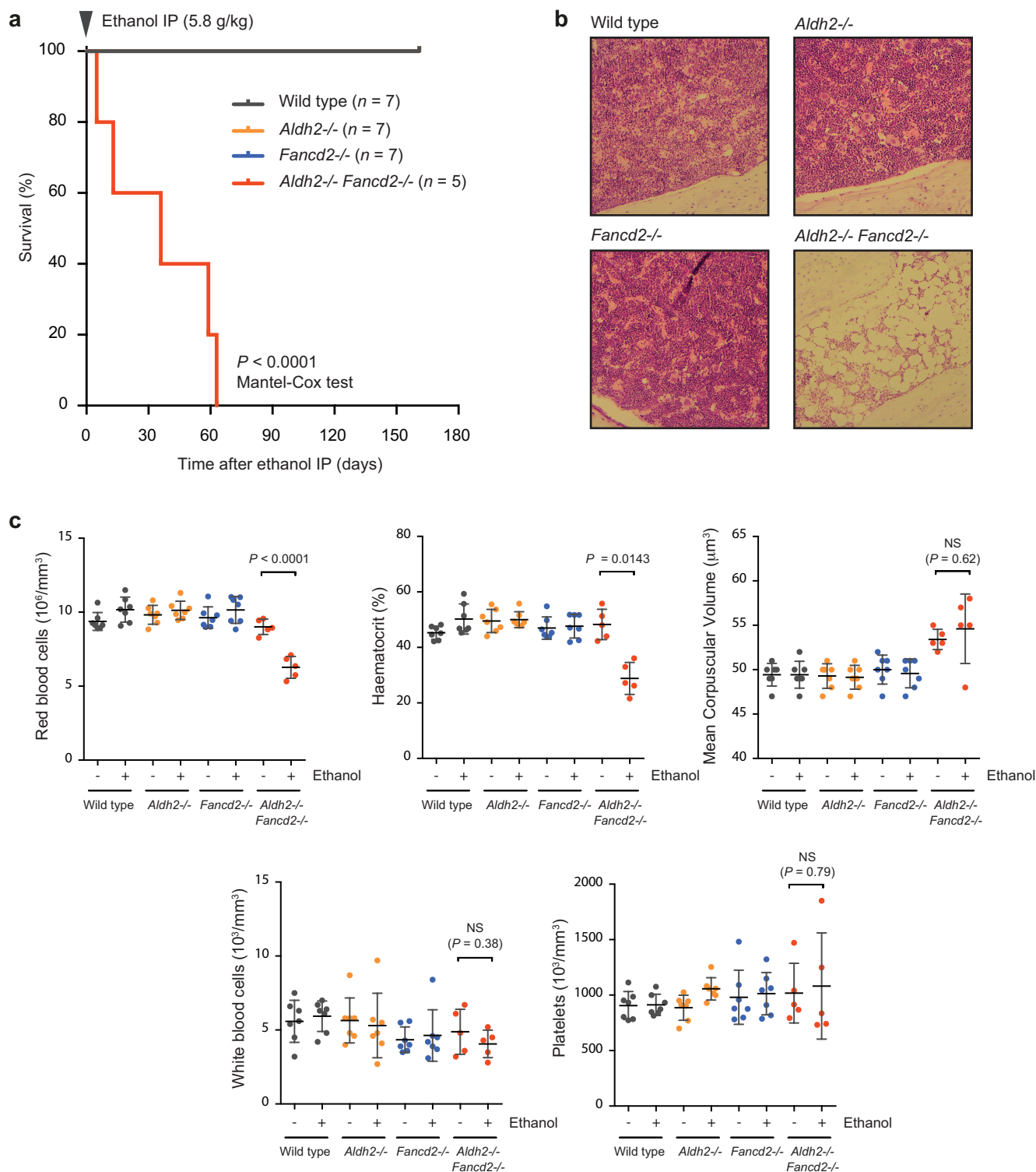


Extended Data Figure 1 | See next page for caption.

Extended Data Figure 1 | Ethanol-induced genomic instability.

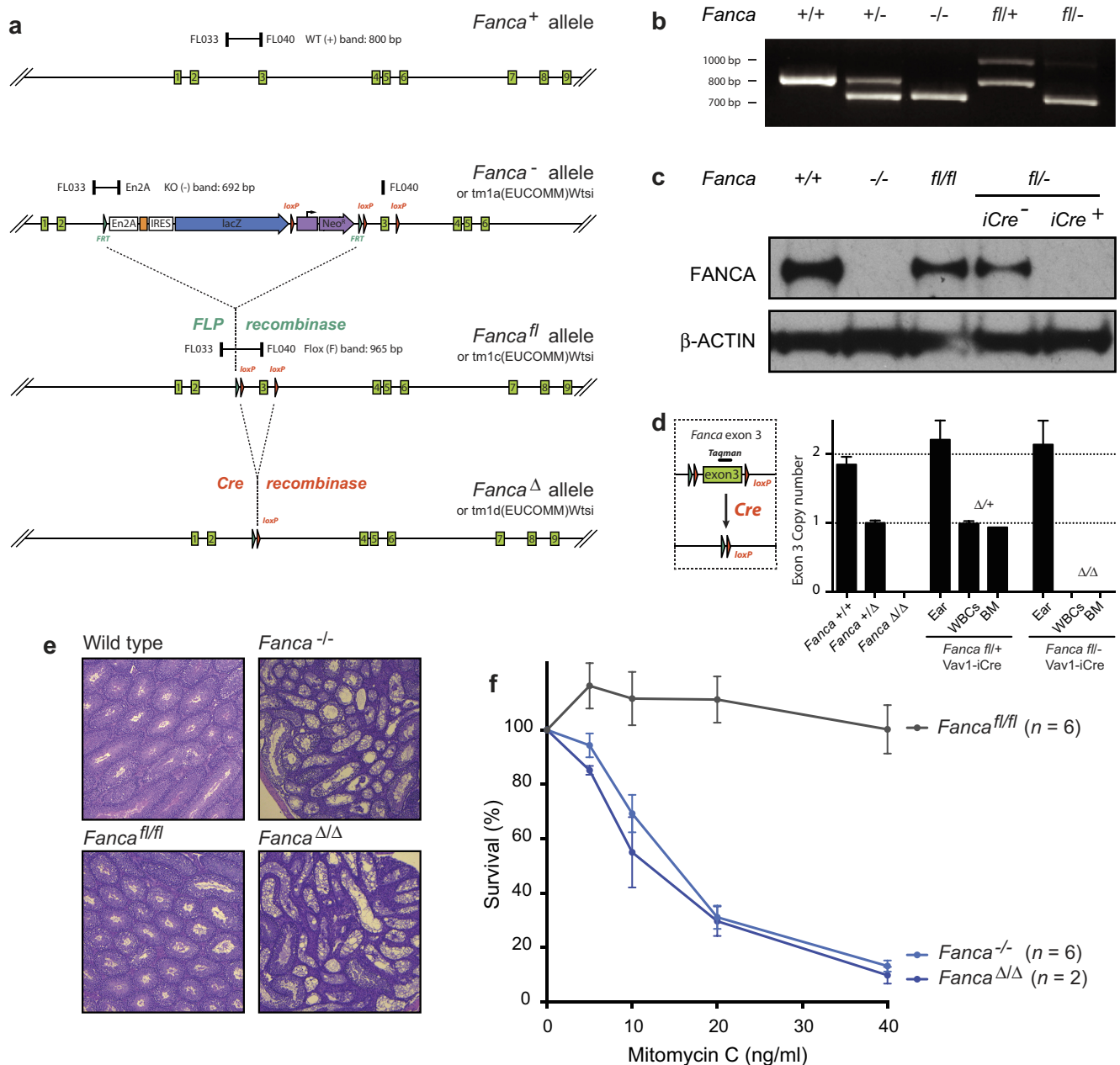
a, Left, representative images of bone marrow metaphase spreads from wild-type mice treated with mitomycin C (MMC); n shows the number of SCE events per metaphase. Right, comparison between number of SCEs in the bone marrow of wild-type and *Aldh2*^{-/-} mice treated with ethanol (5.8 g kg⁻¹) or MMC (1 mg kg⁻¹). Triplicate experiments, 25 metaphases per mouse, $n = 75$; P calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m. Ethanol causes a strong homologous recombination response in *Aldh2*^{-/-} mice, comparable to that observed in wild-type mice exposed to MMC. **b**, Left, representative images of bone marrow metaphase spreads from wild-type and *Fanca*^{-/-} mice; n shows the number of SCE events per metaphase. Right, quantification of SCEs (duplicate experiments, 25 metaphases per mouse, $n = 50$; P calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.). Mice deficient in cross-link repair (*Fanca*^{-/-}, or *Fancd2*^{-/-} in Fig. 1a) show a small but significant increase in the number of spontaneous SCE events, indicating that a homologous recombination repair response occurs in the absence of the Fanconi anaemia pathway. **c**, Scheme depicting the formation of micronucleated erythrocytes. Micronuclei (Mn) generated by fragmentation or mis-segregation of chromosomes during erythrocyte maturation remain in the erythrocyte after extrusion of the main nucleus. These fragments can be detected by

a DNA stain (PI⁺). During maturation, red-cell progenitors lose CD71 expression. Therefore, peripheral CD71⁺ red cells represent immature, short-lived reticulocytes (Ret) and CD71⁻ cells represent mature, long-lived normochromic erythrocytes (NCEs). **d**, Proof-of-principle experiment showing the induction of micronucleated reticulocytes 48 h after MMC treatment (1 mg kg⁻¹). P calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; $n = 29$, 8, 20 and 9 mice, left to right. **e**, Treatment of *Aldh2*^{-/-} mice with ethanol (5.8 g kg⁻¹) leads to potent micronucleus formation. This induction is comparable to that observed in wild-type mice that were treated with the aneugen vincristine (Vcn, 0.2 mg kg⁻¹, 48 h) or clastogenic γ -irradiation (IR, 400 rad, 48 h)⁴⁶. P calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; $n = 29$, 15, 10, 11, 25 and 15 mice. **f**, List of chromosomal aberrations observed in the bone marrow of 8-to-12-week-old untreated *Aldh2*^{-/-}*Fancd2*^{-/-} and control mice. **g**, List of chromosomal aberrations observed in the bone marrow of 8-to-12-week-old *Aldh2*^{-/-}*Fancd2*^{-/-} and control mice 48 h after ethanol treatment (5.8 g kg⁻¹, injected intraperitoneally, IP). In **f** and **g**, three mice and 30 metaphases per mouse were analysed per condition, and the numbers represent the fraction of abnormal metaphases per mouse. **h**, Bar chart classifying the type of aberrations for each genotype (90 metaphases per condition). **i**, Examples of different types of chromosomal aberrations.



Extended Data Figure 2 | A single dose of ethanol precipitates bone-marrow failure in $Aldh2^{-/-} Fancd2^{-/-}$ mice. **a**, A single dose of ethanol (5.8 g kg^{-1} , injected intraperitoneally) leads to anaemia in $Aldh2^{-/-} Fancd2^{-/-}$ mice one to two months after treatment (P calculated by Mantel-Cox test; n = number of mice). **b**, Haematoxylin and eosin

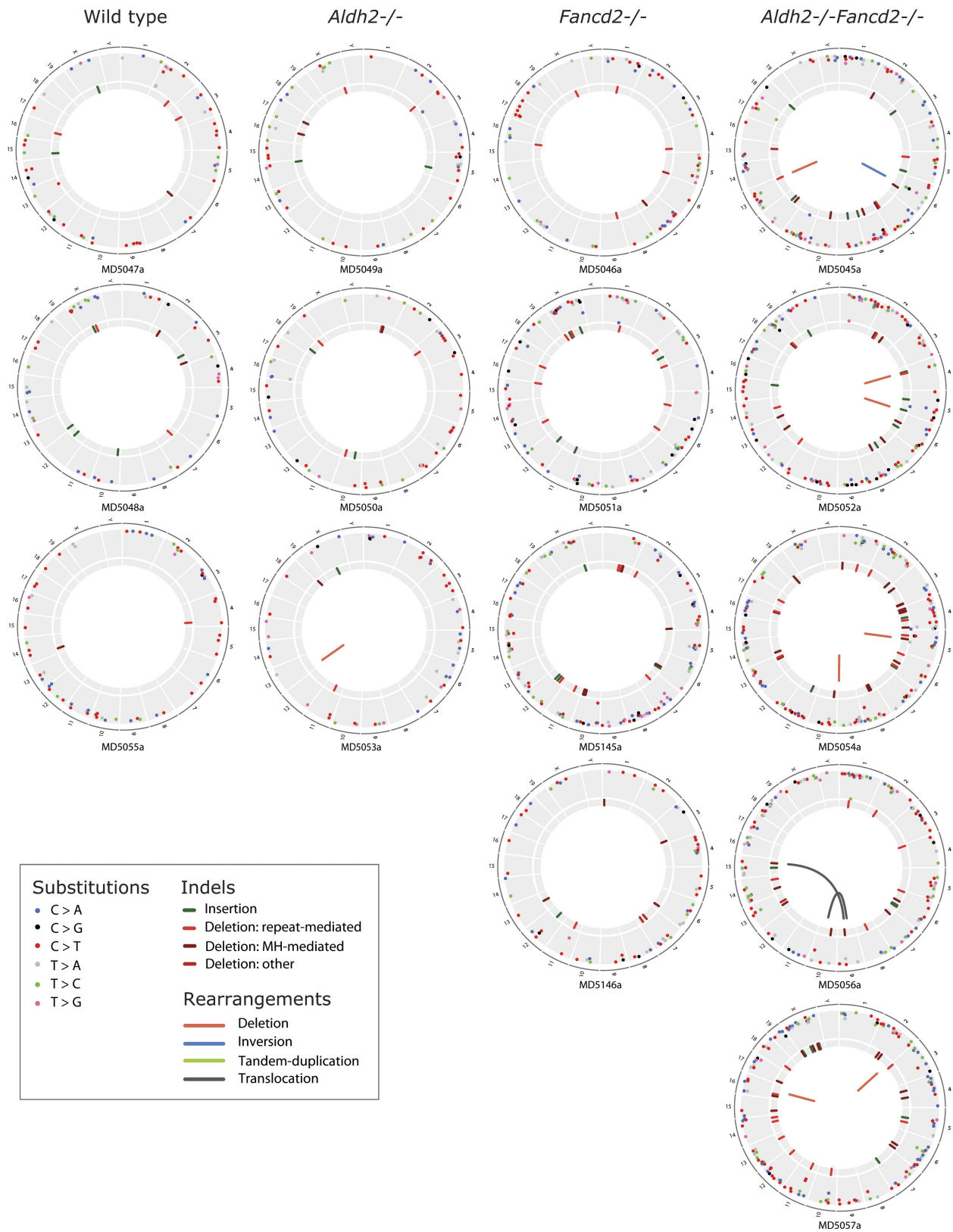
staining of bone marrow sections 30 days after ethanol treatment (original magnification, $\times 100$). **c**, Full blood-count analysis for $Aldh2^{-/-} Fancd2^{-/-}$ and control mice, before injection and terminal bleeds after ethanol treatment (P calculated by paired t -test; data shown as mean and s.e.m.; n = number of mice, as in **a**).



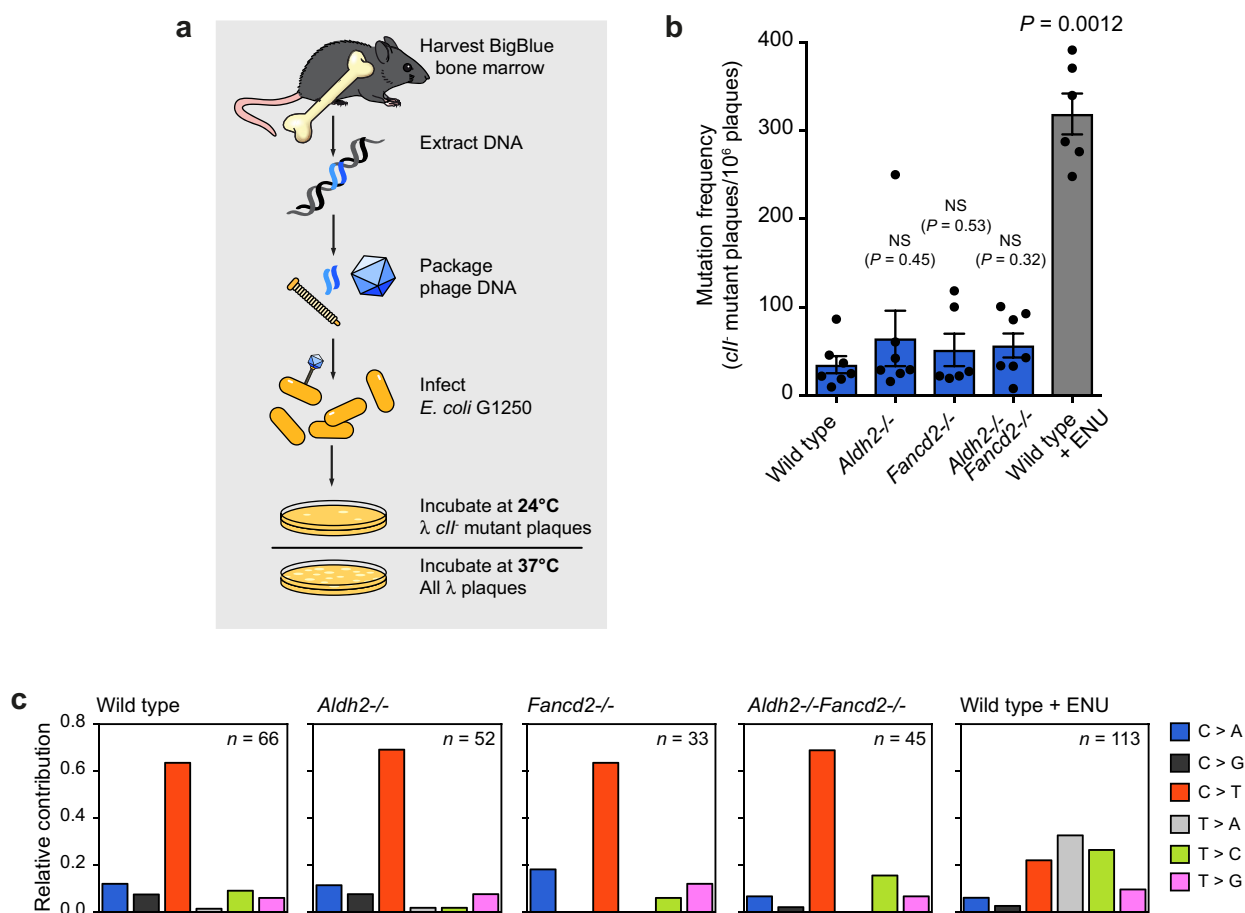
Extended Data Figure 3 | Generation of a conditional *Fanca* allele.

a, Mice carrying the previously reported *Fanca*⁻ allele (*Fanca*^{tm1a(EUCOMM)Wtsi}) were crossed with mice carrying the FLP recombinase, yielding the *Fanca*^{fl} allele (*Fanca*^{tm1c(EUCOMM)Wtsi}). This allele restores FANCA expression as shown by western blot (Fig. 3). Cre-mediated recombination of *Fanca*^{fl} yields the *Fanca*^Δ allele (*Fanca*^{tm1d(EUCOMM)Wtsi}), which lacks exon 3 and leads to loss of FANCA protein (Fig. 3). **b**, Genotyping PCRs for the wild-type, *Fanca*⁻ and *Fanca*^{fl} alleles with primers FL033, FL040 and En2A; showing bands of the expected sizes. **c**, Western blot (single experiment) showing complete absence of FANCA protein in the spleens of *Fanca*^{-/-} and *Fanca*^{fl/-} Vav1-iCre mice. For gel source data, see Supplementary Fig. 1. **d**, Determination of the number of exon 3 copies by quantitative

PCR. Wild-type, *Fanca*^{+/-} and *Fanca*^{Δ/Δ} mice carry 2, 1 and 0 copies, respectively. *Fanca*^{fl} Vav1-iCre mice show tissue-specific deletion of exon 3 in white blood cells (WBCs) and bone marrow (n = 4 technical replicates; bars: mean, s.d.). **e**, Microscopic analysis of haematoxylin and eosin-stained sections of testes (original magnification, ×50) from wild-type, *Fanca*^{-/-}, *Fanca*^{fl/fl} and *Fanca*^{Δ/Δ} males at 12 weeks, showing impaired spermatogenesis in testes of *Fanca*^{-/-} and *Fanca*^{Δ/Δ} mice (one experiment). **f**, Sensitivity assay of transformed mouse-embryonic fibroblasts (MEFs) derived from *Fanca*^{-/-}, *Fanca*^{fl/fl} and *Fanca*^{Δ/Δ} embryos, showing hypersensitivity of both *Fanca*^{-/-} and *Fanca*^{Δ/Δ} cells to the cross-linking agent mitomycin C (n = number of experiments, each carried out in quadruplicate; bars: mean, s.e.m.).

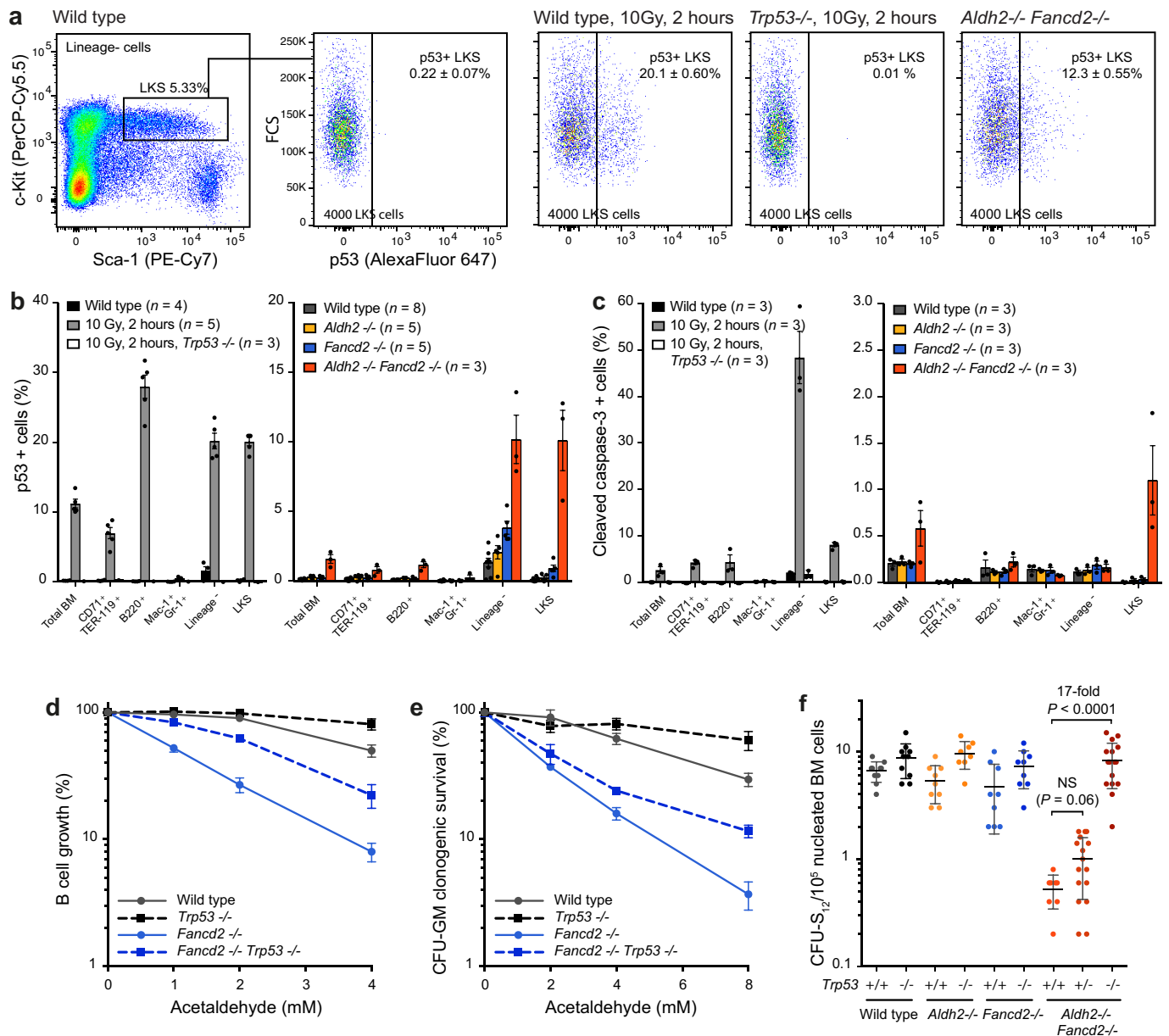


Extended Data Figure 4 | Endogenous aldehydes mutate the HSC genome. Circos plots showing the mutations observed in all sequenced HSC clones (wild type, $n = 3$; *Aldh2*^{-/-}, $n = 3$; *Fancd2*^{-/-}, $n = 4$; and *Aldh2*^{-/-}*Fancd2*^{-/-}, $n = 5$ HSC genomes). Substitutions, indels and rearrangements are plotted.



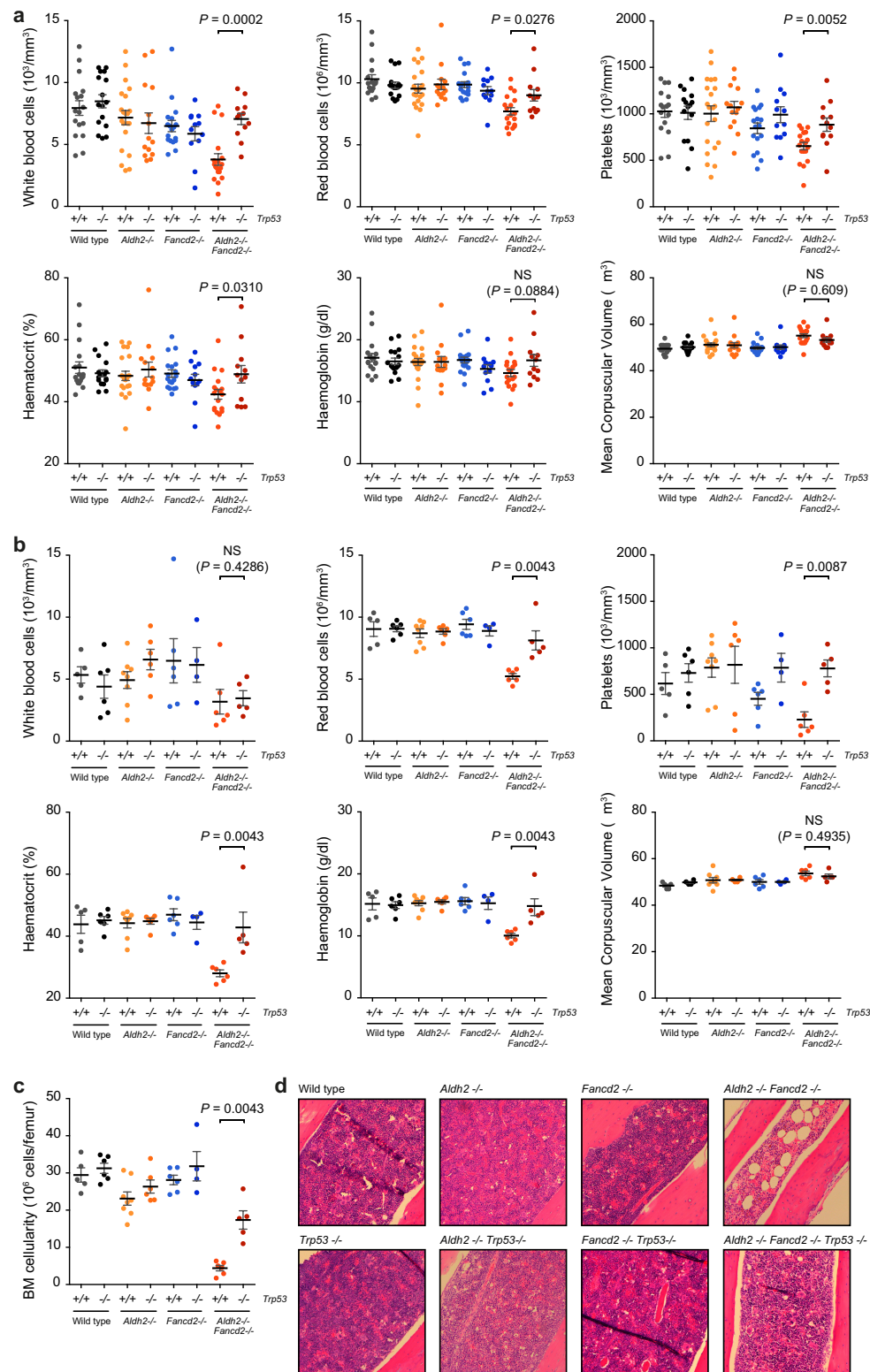
Extended Data Figure 5 | Detection of point mutations in mice with the BigBlue reporter system. **a**, Chromosome 4 of the BigBlue reporter mouse harbours a λ -phage transgene that contains the mutational target. The phage DNA can be recovered from mouse tissues, packaged into phage and used to infect bacteria. Phage *cII*⁻ mutants can be detected by the ability of these phage to form plaques at 24°C. **b**, Quantification of the frequency of *cII*⁻ mutant phage recovered from the bone marrow of young *Aldh2*^{-/-} *Fancd2*^{-/-} and control mice carrying the BigBlue transgene.

ENU-treated mice serve as positive controls for the assay (*P* calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; *n* = 7, 7, 6, 7 and 6 mice, left to right). **c**, Relative contribution of the indicated mutation classes to the point-mutation spectra of *cII*⁻ mutant phage isolated from the bone marrow. The ENU-mutation spectrum is characterized by T to A transversions and T to C transitions. *n* is the number of sequenced *cII*⁻ mutant phage.



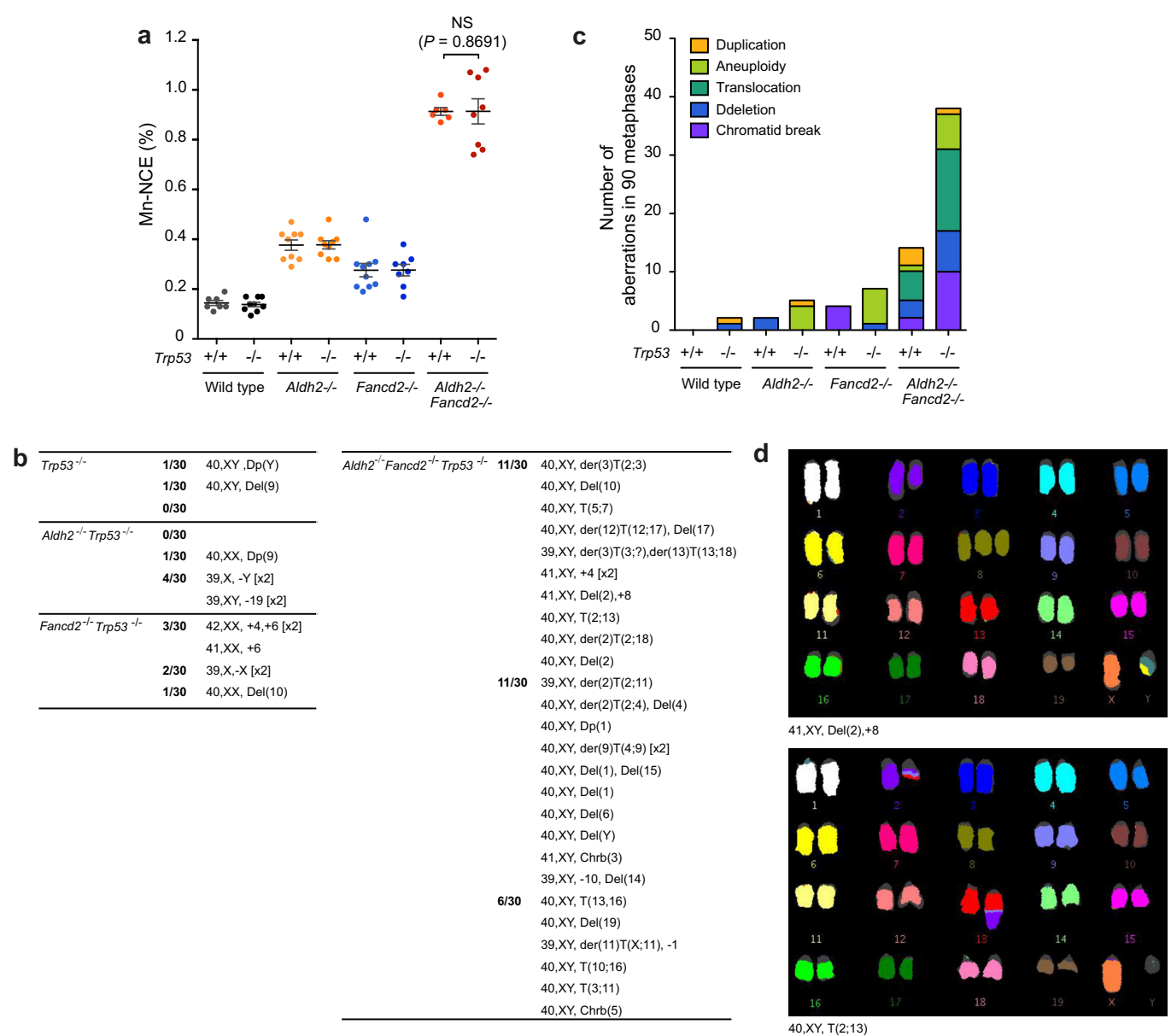
Extended Data Figure 6 | Aldehyde-induced stress elicits a p53 response. **a**, Representative flow cytometry plots for the quantification of p53⁺ LKS cells from 8-to-12-week-old *Aldh2*^{-/-}*Fancd2*^{-/-} and control mice. Cells were collected from wild-type and *Trp53*^{-/-} mice 2 h after 10 Gy irradiation as positive and negative controls, respectively, for the assay. **b**, Quantification of the frequency of p53⁺ cells in different bone-marrow populations. **c**, Quantification of the frequency of cleaved-caspase-3⁺ cells in different bone marrow populations by flow cytometry. In **b** and **c**, irradiated wild-type and *Trp53*^{-/-} mice were used as controls. Owing to the low numbers of LKS CD48⁻ CD150⁺ cells in *Aldh2*^{-/-}*Fancd2*^{-/-} mice, the number of p53⁺ or cleaved-caspase-3⁺ HSCs

could not be determined (data shown as mean and s.e.m.; n = number of mice). **d**, **e**, Survival of B cells and myeloid progenitors (CFU-GM) following exposure to acetaldehyde *in vitro*. Cells were obtained from *Fancd2*^{-/-}*Trp53*^{-/-} and control mice. Each point represents the mean of three independent experiments, each carried out in quadruplicate; data shown as mean and s.e.m. **f**, Frequency of CFU-S₁₂ in the bone marrow of *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} and control mice. Each point represents the number of CFU-S₁₂ in the spleen of a single recipient (P calculated by two-sided Mann-Whitney test; data shown as mean and s.e.m.; n = 10–15 mice).



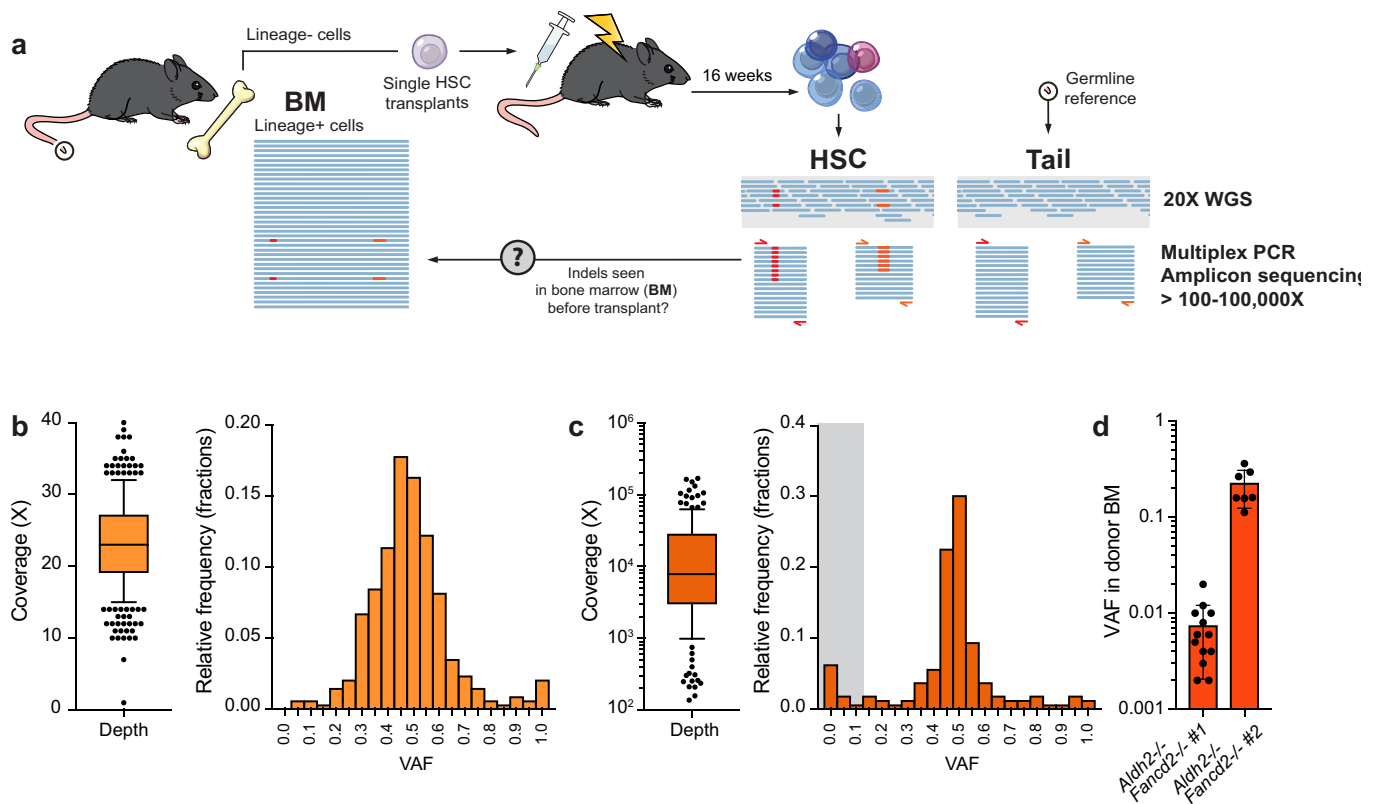
Extended Data Figure 7 | p53 deficiency suppresses peripheral-blood cytopenias and ethanol-induced bone-marrow failure in *Aldh2*^{-/-}*Fancd2*^{-/-} mice. **a**, Full blood count analysis of *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} and control mice (8-to-12 weeks old, on a C57BL/6 × 129S4S6/Sv F1 background). A significant increase in the number of white blood cells, red blood cells, platelets and haematocrit was observed in *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} mice compared to *Aldh2*^{-/-}*Fancd2*^{-/-} mice (*P* calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; *n* = 17, 16, 21, 14, 18, 12, 18 and 12 mice,

left to right). **b**, *Aldh2*^{-/-}*Fancd2*^{-/-}, *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} and control mice were treated with ethanol in their drinking water for 10 days as described previously⁶. Full blood-count analyses were carried out after 10 days of ethanol treatment. **c**, Bone marrow cellularity after 10 days of ethanol treatment. In **b**, **c**, *P* calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; *n* = 5, 6, 8, 6, 4, 6 and 5 mice, left to right). **d**, Haematoxylin and eosin staining of bone-marrow sections 10 days after ethanol treatment (original magnification, ×100).



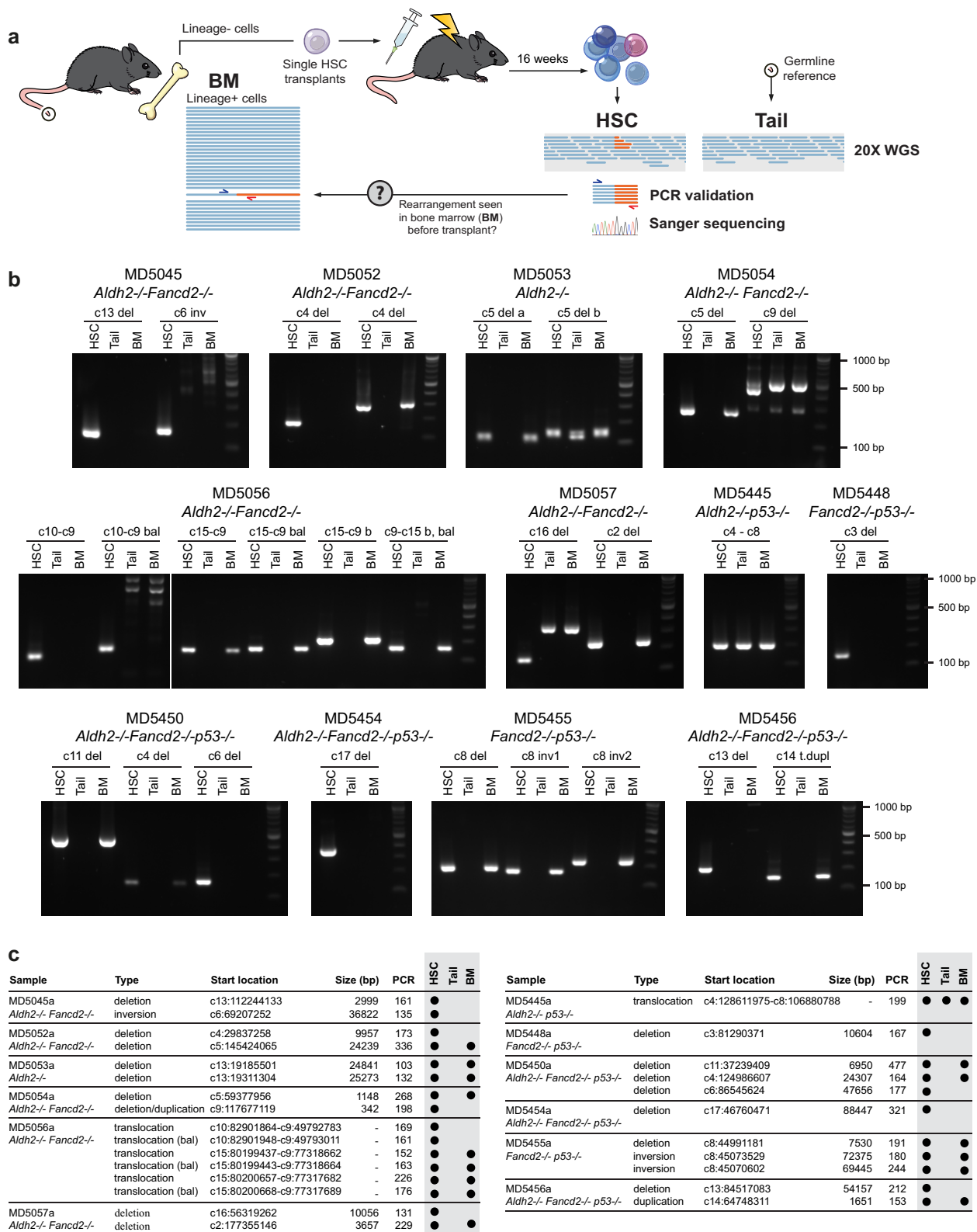
Extended Data Figure 8 | Genomic instability in *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} mice. **a**, Quantification of micronucleated NCEs in the blood of *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} and control mice (*P* calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; *n* = 8 mice). **b**, List of chromosomal aberrations observed in the bone marrow of 8-to-12 week-old untreated *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} and control

mice. Three mice and 30 metaphases per mouse were analysed per genotype; the numbers represent the fraction of abnormal metaphases per mouse. **c**, Bar chart classifying the types of aberrations for each genotype (90 metaphases per condition). **d**, Examples of two metaphases from an *Aldh2*^{-/-}*Fancd2*^{-/-}*Trp53*^{-/-} mouse.



Extended Data Figure 9 | Validation of indels by targeted deep sequencing. **a**, Scheme depicting the generation of HSC clones by transplantation of single stem cells, subsequent whole-genome sequencing and validation of indel calls by amplicon deep sequencing. On the basis of the indel location from 20× whole-genome sequencing, we designed multiplex PCRs and deep sequenced the PCR products to higher coverage (100–100,000×) to confirm that the calls were not sequencing artefacts. In addition, we attempted to detect indels in DNA samples of bone-marrow cells from the mice that provided the transplanted HSCs. **b**, Coverage depth and VAF of the filtered set of indel calls from whole-genome sequencing ($n = 342$ indels; box plot shows the mean, box edges represent the first and third quartiles, whiskers extend over 10–90% of data). **c**, Coverage depth and VAF of the indel calls from deep sequencing validation ($n = 159$ locations; box plot shows the mean, box edges

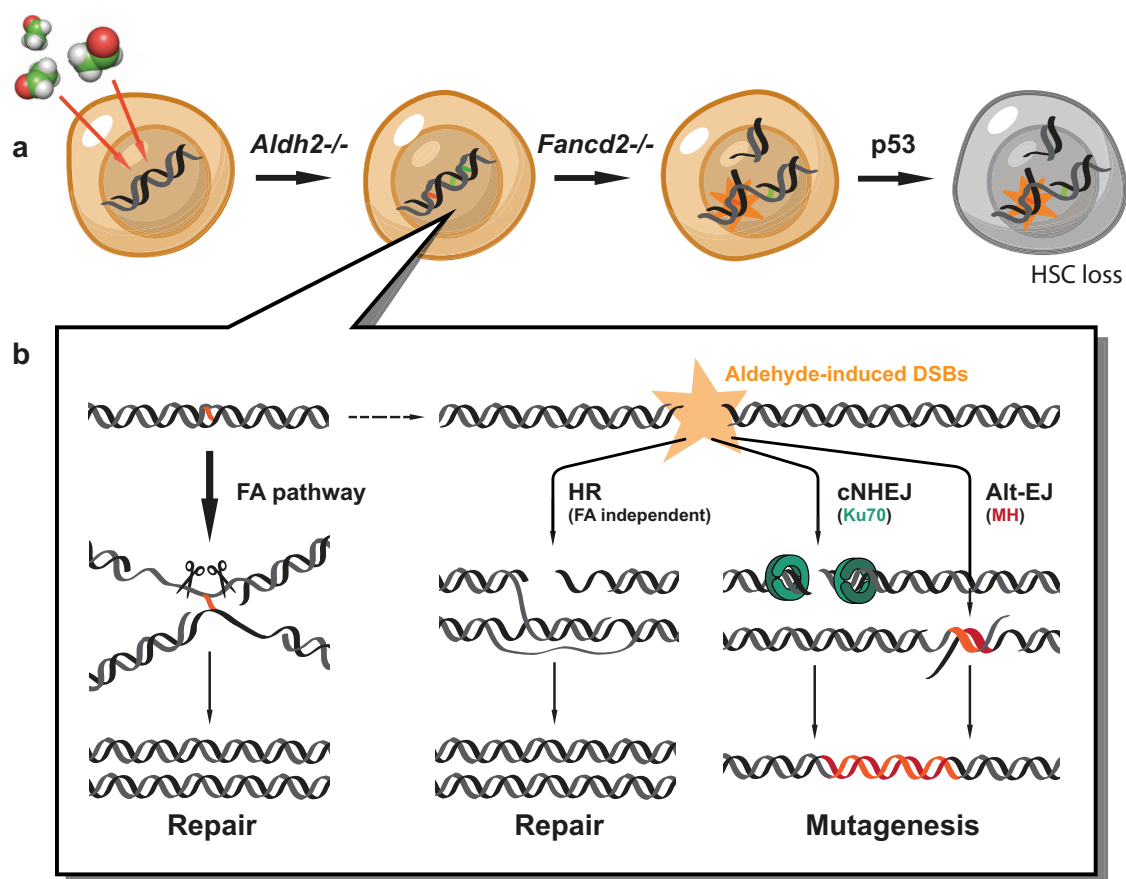
represent the first and third quartiles, whiskers extend over 10–90% of data). One hundred and fifty-nine locations had coverage greater than 100× and were used for the analysis. We could validate the presence of 91.2% of the initial calls; 14/159 (8.8%) calls had VAF < 0.1 and were deemed false positives (indicated by grey shading). Note that the VAF distribution is centred tightly around 0.5, confirming the clonal nature of most indels. **d**, We used targeted deep sequencing to look for indel calls in bone-marrow samples from the mice that provided the transplanted HSCs. In most cases, the calls were below the detection limit of the assay (VAF < 0.0001). However, we could detect indels from two *Aldh2*^{−/−} *Fancd2*^{−/−} HSCs, indicative of ‘clonal haematopoiesis’ in these mice (accounting for 0.7 and 21.4% of blood production, respectively). Data shown as mean and s.e.m.; $n = 13$ and 7 indels.



Extended Data Figure 10 | Validation of rearrangements by PCR.

a, Scheme depicting the generation of HSC clones by transplantation of single stem cells, subsequent whole-genome sequencing and validation of rearrangement calls by PCR. We designed primers for nested PCRs flanking the breakpoints calculated by the BRASS algorithm, and the identity of the products was confirmed by Sanger sequencing. In addition, we attempted to detect the rearrangements in DNA samples of bone-marrow cells from the mice that provided the transplanted HSCs, demonstrating that these changes did not arise during clonal expansion and were present in the stem cell at the time of transplantation. **b**, Agarose

gels (one experiment) showing presence of specific PCR amplification from DNA of HSC clones, absence in matched germline samples from the tail of the same mouse and, in some cases, detection in bone-marrow tissue that predates the transplants. PCR amplification in these samples is dependent on the contribution of the transplanted HSC to blood production, and the sensitivity of each PCR. Gel source data is shown in Supplementary Fig. 1. **c**, List summarizing the rearrangements found in 28 HSC clones and the results from **b**. All 27 rearrangements could be detected by PCR and confirmed by Sanger sequencing, 16/27 (59%) rearrangements could be detected before transplantation.



Extended Data Figure 11 | Mechanisms to maintain genetic integrity and suppress mutagenesis by endogenous aldehydes in HSCs.

a, Aldehyde catabolism and Fanconi anaemia (FA)-pathway-mediated DNA repair constitute two distinct tiers of protection against aldehyde damage. Loss of this protection leads to the accumulation of DNA damage and mutagenesis. Passage of mutated genetic information is prevented by

the activation of p53, leading to HSC loss. **b**, In the absence of a functional Fanconi anaemia pathway, aldehyde lesions degenerate into DNA DSBs that can be repaired through error-free recombination. However, this mechanism is not sufficient to fully compensate for Fanconi anaemia inactivation, leading to the engagement of both classical and alternative end-joining, and subsequent mutagenic repair.

Rotation in [C II]–emitting gas in two galaxies at a redshift of 6.8

Renske Smit^{1,2}, Rychard J. Bouwens³, Stefano Carniani^{1,2}, Pascal A. Oesch⁴, Ivo Labbé³, Garth D. Illingworth⁵, Paul van der Werf³, Larry D. Bradley⁶, Valentino Gonzalez^{7,8}, Jacqueline A. Hodge³, Benne W. Holwerda⁹, Roberto Maiolino^{1,2} & Wei Zheng¹⁰

The earliest galaxies are thought to have emerged during the first billion years of cosmic history, initiating the ionization of the neutral hydrogen that pervaded the Universe at this time. Studying this ‘epoch of reionization’ involves looking for the spectral signatures of ancient galaxies that are, owing to the expansion of the Universe, now very distant from Earth and therefore exhibit large redshifts. However, finding these spectral fingerprints is challenging. One spectral characteristic of ancient and distant galaxies is strong hydrogen-emission lines (known as Lyman- α lines), but the neutral intergalactic medium that was present early in the epoch of reionization scatters such Lyman- α photons. Another potential spectral identifier is the line at wavelength 157.4 micrometres of the singly ionized state of carbon (the [C II] $\lambda = 157.74\ \mu\text{m}$ line), which signifies cooling gas and is expected to have been bright in the early Universe. However, so far Lyman- α -emitting galaxies from the epoch of reionization have demonstrated much fainter [C II] luminosities than would be expected from local scaling relations^{1–5}, and searches for the [C II] line in sources without Lyman- α emission but with photometric redshifts greater than 6 (corresponding to the first billion years of the Universe) have been unsuccessful. Here we identify [C II] $\lambda = 157.74\ \mu\text{m}$ emission from two sources that we selected as high-redshift candidates on the basis of near-infrared photometry; we confirm that these sources are two galaxies at redshifts of $z = 6.8540 \pm 0.0003$ and $z = 6.8076 \pm 0.0002$. Notably, the luminosity of the [C II] line from these galaxies is higher than that found previously in star-forming galaxies with redshifts greater than 6.5. The luminous and extended [C II] lines reveal clear velocity gradients that, if interpreted as rotation, would indicate that these galaxies have similar dynamic properties to the turbulent yet rotation-dominated disks that have been observed in H α -emitting galaxies two billion years later, at ‘cosmic noon’.

Using the Atacama Large Millimetre Array (ALMA) in Chile, we obtained spectroscopy at 241–245 GHz for two Lyman-break galaxies (LBGs)—COS-3018555981 and COS-2987030247—at an estimated photometric redshift of just less than 7, corresponding to roughly 800 million years after the Big Bang. These two sources are luminous in the rest-frame ultraviolet (UV; L_{UV} is roughly $2 \times L_{z=7}^*$ (this latter value being obtained from ref. 6)), but are representative of ‘normal’ star-forming galaxies at redshifts of around 7, with a UV-based star-formation rate (SFR) of $(19\text{--}23)M_{\odot}\ \text{yr}^{-1}$ (L , luminosity; M_{\odot} , mass of the Sun). We selected these sources on the basis of the blue rest-frame optical colours measured in the 3.6- μm and 4.5- μm photometric bands by the Spitzer Space Telescope⁷; these rest-frame colours strongly constrain the photometric redshift probability distribution to the range $6.6 < z < 6.9$. The two sources are among the most extreme [O III] + H β

emitters known at redshifts of around 7 (refs 7, 8). We observed them using a 36-antennae ALMA configuration (angular resolution $1.1'' \times 0.7''$, equivalent to $5.8\ \text{kpc} \times 3.7\ \text{kpc}$ at $z = 6.8$), with 24 minutes of source-integration time for each target. Using this spectral scan, we searched for [C II] lines in the redshift range $z_{[\text{C II}]} = 6.74\text{--}6.90$. Our results are summarized in Extended Data Table 1.

We detected a line at $241.97 \pm 0.01\ \text{GHz}$ and at $243.42 \pm 0.01\ \text{GHz}$ for COS-3018555981 and COS-2987030247, respectively, in both one-dimensional spectra and spectral-line-averaged maps (Fig. 1; more than 5σ significance). We thereby derived spectroscopic redshifts of $z_{[\text{C II}]} = 6.8540 \pm 0.0003$ and $z_{[\text{C II}]} = 6.8076 \pm 0.0002$, respectively, in excellent agreement with the photometric redshift estimates of 6.76 ± 0.07 and 6.66 ± 0.14 for COS-3018555981 and COS-2987030247; we also derived line-widths of $232 \pm 30\ \text{km s}^{-1}$ and $124 \pm 18\ \text{km s}^{-1}$, respectively. Although successful line searches have previously confirmed far-infrared lines in submillimetre-selected star-bursting galaxies at redshifts of more than 6 (refs 9, 10), and a few tantalizing ‘blind’ candidate [C II] emitters (with no optical or near-infrared counterpart) have been detected with a significance of around 4σ (ref. 11), this is the first time that normal star-forming galaxies in this early epoch—selected at optical or near-infrared wavelengths—have confidently been spectroscopically confirmed with ALMA.

We furthermore obtained upper limits to the far-infrared dust-continuum emission from the ALMA data. We found infrared SFRs of less than $(16\text{--}19)M_{\odot}\ \text{yr}^{-1}$ —rates that are consistent with ‘normal’ star-forming galaxies in the local Universe¹², and which rule out the presence of a dusty starburst in these sources. Figure 2 shows that, given the colour of the UV-continuum slopes of these sources (this slope, β_{UV} , is roughly -1.2), a higher dust content (measured by the infrared excess, $L_{\text{IR}}/L_{\text{UV}}$) would be expected if the sources obeyed the Meurer dust law, which is observed to apply for local starburst galaxies¹³. Scatter in the IRX- β_{UV} relation could be due to the geometry of the dust, the age of the population of galaxies, or the shape of the attenuation curve. However, for blue galaxies (where β_{UV} is less than around -0.5) that scatter below the Meurer¹³ relation—such as our selected galaxies—the most likely way to reproduce the low observed values of IRX is through a steeper attenuation curve, such as that derived for the Small Magellanic Cloud¹⁴ (consistent with our measurements to within 3σ), in combination with a potential increase in dust temperature at higher redshifts.

In Fig. 2 we present the measured flux of the [C II] lines as a function of the SFR, which is consistent with the SFR- $L_{[\text{C II}]}$ relation for galaxies in the local Universe (ref. 15), and consistent with data for similarly bright galaxies observed at redshifts of about 5–6 (refs 16, 17). By contrast, [C II] observations from the epoch of reionization so far have shown

¹Cavendish Laboratory, University of Cambridge, 19 JJ Thomson Avenue, Cambridge CB3 0HE, UK. ²Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK. ³Leiden Observatory, Leiden University, NL-2300 RA Leiden, Netherlands. ⁴Observatoire de Genève, Université de Genève, Chemin des Maillettes 51, CH-1290 Versoix, Switzerland. ⁵UCO/Lick Observatory, University of California, Santa Cruz, 1156 High Street, Santa Cruz, California 95064, USA. ⁶Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. ⁷Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago 7591245, Chile. ⁸Centro de Astrofísica y Tecnologías Afines (CATA), Camino del Observatorio 1515, Las Condes, Santiago 7591245, Chile. ⁹Department of Physics and Astronomy, University of Louisville, Louisville, Kentucky 40292, USA. ¹⁰Department of Physics and Astronomy, The Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA.

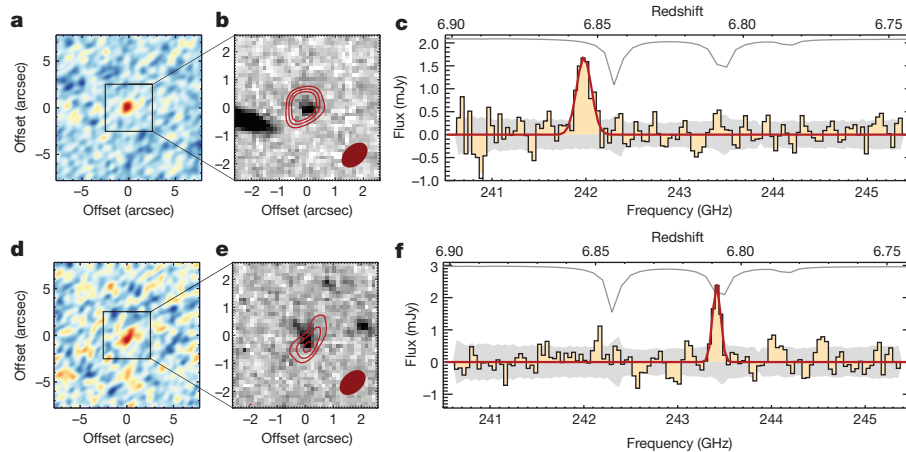


Figure 1 | Spectroscopic line confirmations of the galaxies targeted in this study. ALMA line maps and spectra for two galaxies with photometric redshifts (z_{phot}) in the range $6.6 < z_{\text{phot}} < 6.9$ (ref. 7). We detect a 8.2σ [C II] line at $z_{\text{[C II]}} = 6.8540 \pm 0.0003$ in galaxy COS-3018555981 (a–c), and a roughly 5.1σ [C II] line at $z_{\text{[C II]}} = 6.8076 \pm 0.0002$ in galaxy COS-2987030247 (d–f). a, d, $20'' \times 20''$ images of the ALMA cube (before primary-beam correction), collapsed over 241.85–242.10 GHz for COS-3018555981 and 243.35–243.45 GHz for COS-2987030247 (with root-mean-squares of 0.1 mJy and 0.2 mJy, respectively). b, e, $5'' \times 5''$ images

of the targeted sources. Hubble Space Telescope H_{160} -band imaging is shown in greyscale; the overlaid red contours show the 3σ , 4σ and 5σ levels of the spectral-line-averaged maps on the left. The filled ellipses in the bottom right corners indicate the beam size ($1.1'' \times 0.7''$ half-power widths). c, f, The spectra extracted from within a contour of the half-maximum power in the line maps. The red lines show the best-fitting Gaussian line profiles; the grey lines at the top show the atmospheric absorption; the grey filled regions give the $\pm 1\sigma$ noise in the spectrum.

that these galaxies fall substantially below the local relation^{1–5}. This is probably because we chose our $z > 6.5$ targets differently to previous authors: we selected [O III] + H β emitters as opposed to Lyman- α -emitting galaxies.

Our sources have slightly higher SFRs and redder UV slopes (at roughly -1.2) than previously studied galaxies from this epoch, which could indicate that our galaxies are more evolved and more metal rich. Sources with extremely low oxygen abundance in the local Universe are typically found to be [C II] deficient^{15,18} owing to their hard radiation field, and therefore metallicity could be an important discriminator between [C II]-bright and [C II]-faint sources¹⁹. Moreover, in local galaxies the SFR surface density (Σ_{SFR}) drives a continuous trend of deepening [C II] deficit as a function of increasing Σ_{SFR} (refs 18, 20), indicating that local processes such as the radiation-field intensity are important in driving [C II] luminosity. If [C II]-faint sources at $z > 6$, currently unresolved in [C II] lines, have higher star-formation surface brightness than our galaxies, this could also explain the different SFR/ $L_{\text{[C II]}}$ ratios.

Furthermore, our sources have high-equivalent-width optical emission lines, which could suggest an ongoing starburst and potentially a higher fraction of [C II] emission emerging from H II regions. Starbursts and H II galaxies in the local Universe have slightly elevated [C II] luminosities for a given SFR¹⁵, and therefore we could specifically be targeting the brightest [C II] galaxies of the overall $z \approx 7$ galaxy population. Finally, while we do not have spectroscopy covering the Lyman- α line for COS-3018555981 and COS-2987030247, our sources could be weaker Lyman- α emitters than are typically seen in spectroscopically confirmed sources at this redshift. Lyman- α emission is suggested to be inversely correlated with neutral gas column density²¹ and can therefore affect the visibility of [C II], which emerges both in the diffuse neutral and in the warm ionized medium of a galaxy.

We also determined [C II] half-light radii (deconvolved from the beam size) of 2.6 ± 0.8 kpc and 3.1 ± 1.0 kpc for COS-3018555981 and COS-2987030247—nearly twice the half-light radius of the UV in the brightest LBGs at this redshift²². We used the spatial extent of the [C II] detection to investigate the velocity structure of these sources, which reveals a projected velocity difference over the galaxy of 111 ± 28 km s⁻¹ and 54 ± 20 km s⁻¹ for COS-3018555981 and COS-2987030247, respectively (Fig. 3), similar to the velocity gradients observed recently in two galaxies at redshifts of around 5–6 (refs 23, 24). Given the low angular

resolution of the observations, there are various ways to interpret these velocity gradients. A rotating galaxy disk would be one interpretation; however, a merger involving one or more [C II]-emitting galaxies, smoothed by the beam size, could also appear as a regular rotational field. Furthermore, a bipolar outflow or perhaps an inflow of gas could provide an additional velocity component to the [C II] line that might give the impression of galaxy rotation.

We applied an observational criterion for the classification of rotation- and dispersion-dominated systems, based on the full observed velocity gradient, Δv_{obs} , and the integrated line width, σ_{tot} , of a galaxy, such that $\Delta v_{\text{obs}}/2\sigma_{\text{tot}}$ values of more than 0.4 are likely to be rotation-dominated sources²⁵. In Fig. 4, we compare this quantity for

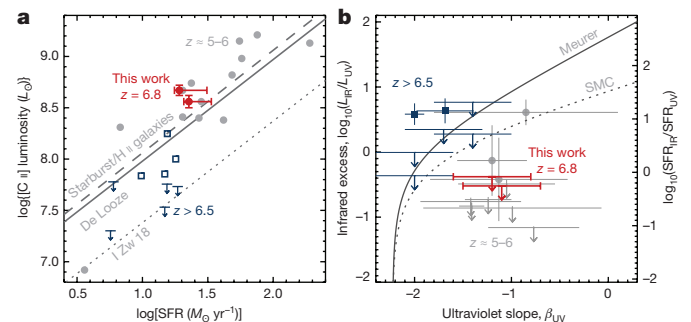


Figure 2 | [C II] luminosity and dust continuum for $z > 5$ galaxies. a, [C II] line luminosity as a function of the SFR for COS-3018555981 and COS-2987030247 (red points; error bars for the SFRs reflect 1σ upper limits on the infrared continuum), compared with [C II] detections at redshifts of around 5–6 (light grey points)^{4,16,17} and more than 6.5 (blue open squares and arrows)^{1–3,5}. Locally observed relations¹⁵ are indicated by solid lines (star-forming galaxies) and dashed lines (starburst and H II galaxies). The dotted line gives the 0.6-dex offset from the local relation found for the dwarf galaxy I Zw 18 ('I Zw 18'). b, Infrared excess ($L_{\text{UV}}/L_{\text{IR}}$) as a function of the UV-continuum slope (β_{UV}) of our sources compared with expectations from the Meurer¹³ relation (solid grey line) and a similar relation based on the dust law of the Small Magellanic Cloud¹⁴ (SMC; dotted grey line). We include [C II] detections at redshifts of around 5–6 as light grey points¹⁶, and detections (and upper limits) at redshifts greater than 6.5 as blue solid squares (and arrows)^{28–30}. Upper limits and error bars represent 1σ significance levels. L_{\odot} , luminosity of the Sun.

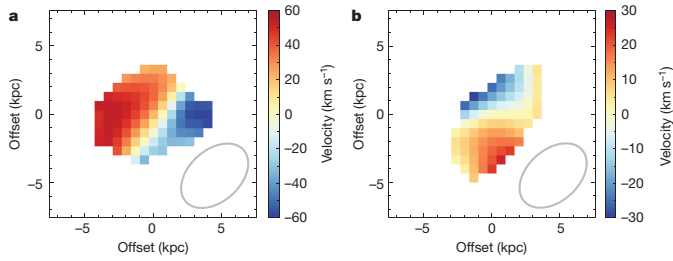


Figure 3 | Velocity structure of the detected [C II] emission in COS-3018555981 and COS-2987030247. **a, b,** Velocity fields measured in COS-3018555981 (**a**) and COS-2987030247 (**b**). The observations are spatially resolved, as shown by the beam size of the observations (grey ellipses), and reveal a projected velocity difference over the galaxies of $111 \pm 28 \text{ km s}^{-1}$ and $54 \pm 20 \text{ km s}^{-1}$, respectively. Given the low angular resolution of the observations, we could interpret the velocity gradients as disk rotation or alternatively perhaps as a merging system with two or more velocity components.

our galaxies with that measured through H α emission for galaxies at redshifts of around 1 to 3 (ref. 25). Although our sources are an order of magnitude smaller in terms of stellar mass, and at an epoch 2.5 billion years earlier in cosmic time, we find $\Delta v_{\text{obs}}/2\sigma_{\text{tot}}$ values of 0.57 ± 0.16 and 0.52 ± 0.21 for COS-3018555981 and COS-2987030247—similar to the values for the turbulent yet rotationally supported galaxy disks at redshifts of about 2 (ref. 25). Assuming a circularly symmetric galaxy disk model, we estimate dynamic masses, M_{dyn} , of $1.0^{+0.3}_{-0.2} \times 10^{10} M_{\odot}$ and $0.4^{+0.9}_{-0.3} \times 10^{10} M_{\odot}$ for COS-3018555981 and COS-2987030247, respectively. (Note, however, that the influence of turbulence in these sources could increase the dynamic mass estimates, although by at most a factor of two.) Therefore, these sources have around four to ten times less mass than the bright, UV-selected sources observed recently at redshifts of around 5 to 6 (corresponding to just 200–300 million years later in cosmic time¹⁶), which otherwise appear similar in their [C II] and infrared properties (Fig. 2). Furthermore, the stellar mass in our sources makes up about 14% and 43% of the total dynamic mass that we measure (Fig. 4), in good

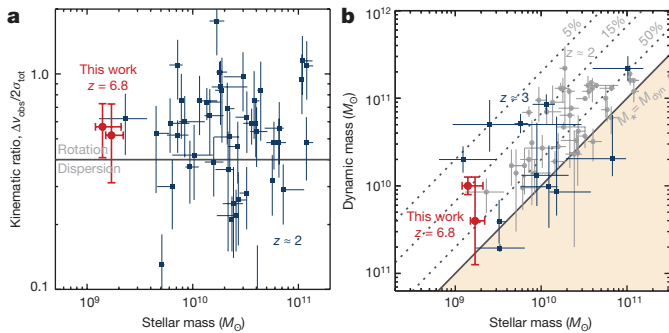


Figure 4 | Dynamic classification and masses of galaxies with redshifts of around 2 or more. **a,** The observed kinematic ratio of the projected velocity range of a galaxy over the velocity dispersion of the system ($\Delta v_{\text{obs}}/2\sigma_{\text{tot}}$) as a function of stellar mass, for COS-3018555981 and COS-2987030247 (red points), and for H α -emitting galaxies from the SINS²⁵ spectroscopy survey at redshifts of about 2 (blue squares). Galaxies with $\Delta v_{\text{obs}}/2\sigma_{\text{tot}}$ ratios of more than 0.4 are classified as probable rotation-dominated systems, while sources with $\Delta v_{\text{obs}}/2\sigma_{\text{tot}}$ ratios of less than 0.4 are probably dispersion-dominated (demarcated by the grey line)²⁵. **b,** Dynamic (total) mass within a roughly 2-kpc half-light radius (assuming a circularly symmetric thin-disk model) is plotted against stellar mass for our sources (red points). Grey dotted lines indicate stellar mass as a fraction of total dynamic mass; the stellar-mass fractions of 14% and 43% for COS-3018555981 and COS-2987030247 are in good agreement with the range of values found for galaxies in the AMAZE survey²⁶ at redshifts of about 3 (blue squares) and in the SINS survey²⁵ for redshifts of about 2 (grey points). Error bars represent 1 σ .

agreement with the 33% stellar mass estimated for the UV-selected sources at redshifts of about 5–6 (ref. 16), and consistent with the wide range of values observed for star-forming galaxies at redshifts of around 1–3 (refs 25, 26). These results indicate a substantial gas fraction in the inner few kiloparsecs of our galaxies, consistent with hydrodynamic simulations of star-forming galaxies at this epoch²⁷.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 June; accepted 3 October 2017.

- Ouchi, M. *et al.* An intensely star-forming galaxy at $z \sim 7$ with low dust and metal content revealed by deep ALMA and HST observations. *Astrophys. J.* **778**, 102 (2013).
- Ota, K. *et al.* ALMA observation of 158 [C II] line and dust continuum of a $z = 7$ normally star-forming galaxy in the epoch of reionization. *Astrophys. J.* **792**, 34 (2014).
- Maiolino, R. *et al.* The assembly of ‘normal’ galaxies at $z \sim 7$ probed by ALMA. *Mon. Not. R. Astron. Soc.* **452**, 54–68 (2015).
- Knudsen, K. K. *et al.* [C II] emission in $z \sim 6$ strongly lensed, star-forming galaxies. *Mon. Not. R. Astron. Soc.* **462**, L6–L10 (2016).
- Pentericci, L. *et al.* Tracing the reionization epoch with ALMA: [C II] emission in $z \sim 7$ galaxies. *Astrophys. J.* **829**, L11 (2016).
- Stark, D. P. Galaxies in the first billion years after the Big Bang. *Annu. Rev. Astron. Astrophys.* **54**, 761–803 (2016).
- Smit, R. *et al.* High-precision photometric redshifts from Spitzer/IRAC: extreme [3.6]–[4.5] colors identify galaxies in the redshift range $z \sim 6.6$ –6.9. *Astrophys. J.* **801**, 122 (2015).
- Smit, R. *et al.* Evidence for ubiquitous high-equivalent-width nebular emission in $z \sim 7$ galaxies: toward a clean measurement of the specific star-formation rate using a sample of bright, magnified galaxies. *Astrophys. J.* **784**, 58 (2014).
- Riechers, D. A. *et al.* A dust-obscured massive maximum-starburst galaxy at a redshift of 6.34. *Nature* **496**, 329–333 (2013).
- Strandet, M. L. *et al.* ISM properties of a massive dusty star-forming galaxy discovered at $z \sim 7$. *Astrophys. J.* **842**, L15 (2017).
- Aravena, M. *et al.* The ALMA spectroscopic survey in the Hubble ultra deep field: search for [C II] line and dust emission in $6 < z < 8$ galaxies. *Astrophys. J.* **833**, 71 (2016).
- Leroy, A. K. *et al.* Estimating the star formation rate at 1 kpc scales in nearby galaxies. *Astron. J.* **144**, 3 (2012).
- Meurer, G. R., Heckman, T. M. & Calzetti, D. Dust absorption and the ultraviolet luminosity density at $z \sim 3$ as calibrated by local starburst galaxies. *Astrophys. J.* **521**, 64–80 (1999).
- Prevot, M. L., Lequeux, J., Prevot, L., Maurice, E. & Rocca-Volmerange, B. The typical interstellar extinction in the Small Magellanic Cloud. *Astron. Astrophys.* **132**, 389–392 (1984).
- De Looze, I. *et al.* The applicability of far-infrared fine-structure lines as star formation rate tracers over wide ranges of metallicities and galaxy types. *Astron. Astrophys.* **568**, A62 (2014).
- Capak, P. L. *et al.* Galaxies at redshifts 5 to 6 with systematically low dust content and high [C II] emission. *Nature* **522**, 455–458 (2015).
- Willott, C. J., Carilli, C. L., Wagg, J. & Wang, R. Star formation and the interstellar medium in $z > 6$ UV-luminous Lyman-break galaxies. *Astrophys. J.* **807**, 180 (2015).
- Smith, J. D. T. *et al.* The spatially resolved [C II] cooling line deficit in galaxies. *Astrophys. J.* **834**, 5 (2017).
- Vallini, L., Ferrara, A., Pallottini, A. & Gallerani, S. Molecular cloud photoevaporation and far-infrared line emission. *Mon. Not. R. Astron. Soc.* **467**, 1300–1312 (2017).
- Herrera-Camus, R. *et al.* [C II] 158 μm emission as a star formation tracer. *Astrophys. J.* **800**, 1 (2015).
- Stark, D. P. *et al.* Ly α and C III] emission in $z = 7$ –9 galaxies: accelerated reionization around luminous star-forming systems? *Mon. Not. R. Astron. Soc.* **464**, 469–479 (2017).
- Bowler, R. A. A., Dunlop, J. S., McLure, R. J. & McLeod, D. J. Unveiling the nature of bright $z \sim 7$ galaxies with the Hubble Space Telescope. *Mon. Not. R. Astron. Soc.* **466**, 3612–3635 (2017).
- Riechers, D. A. *et al.* ALMA imaging of gas and dust in a galaxy protocluster at redshift 5.3: [C II] emission in ‘typical’ galaxies and dusty starbursts ~ 1 billion years after the Big Bang. *Astrophys. J.* **796**, 84 (2014).
- Pavesi, R. *et al.* ALMA reveals weak [N II] emission in ‘typical’ galaxies and intense starbursts at $z = 5$ –6. *Astrophys. J.* **832**, 151 (2016).
- Förster Schreiber, N. M. *et al.* The SINS survey: SINFONI integral field spectroscopy of $z \sim 2$ star-forming galaxies. *Astrophys. J.* **706**, 1364–1428 (2009).
- Gnerucci, A. *et al.* Dynamical properties of AMAZE and LSD galaxies from gas kinematics and the Tully–Fisher relation at $z \sim 3$. *Astron. Astrophys.* **528**, A88 (2011).
- Fiacconi, D. *et al.* Young and turbulent: the early life of massive galaxy progenitors. *Mon. Not. R. Astron. Soc.* **467**, 4080–4100 (2017).

28. Watson, D. *et al.* A dusty, normal galaxy in the epoch of reionization. *Nature* **519**, 327–330 (2015).
29. Schaerer, D. *et al.* New constraints on dust emission and UV attenuation of $z = 6.5$ – 7.5 galaxies from millimeter observations. *Astron. Astrophys.* **574**, A19 (2015).
30. Laporte, N. *et al.* Dust in the reionization era: ALMA observations of a $z = 8.38$ gravitationally lensed galaxy. *Astrophys. J.* **837**, L21 (2017).

Acknowledgements This paper makes use of the following ALMA data: ADS/JAO.ALMA#2015.1.01111.S (http://almascience.org/aq?project_code=2015.1.01111.S). ALMA is a partnership of the European Southern Observatory (ESO, UK), the National Science Foundation (NSF, USA) and the National Institute of Natural Sciences (NINS, Japan), together with the National Research Council (NRC, Canada), the National Security Council (NSC) and Academia Sinica Institute of Astronomy and Astrophysics (ASIAA, Taiwan), and the Korea Astronomy and Space Science Institute (KASI, South Korea), in cooperation with Chile. The Joint ALMA Observatory is operated by ESO, Associated Universities Inc. (AUI)/National Radio Astronomy Observatory (NRAO), and the National Astronomical Observatory of Japan (NAOJ). This work

is part of a Rubicon programme, ‘A multi-wavelength view of the first galaxies’, with project number 680-50-1518, which is financed by the Netherlands Organisation for Scientific Research (NWO). R.M. and S.C. acknowledge ERC Advanced Grant 695671 ‘QUENCH’ and support by the Science and Technology Facilities Council (STFC).

Author Contributions R.S. performed the data reduction and nearly all data analysis for this work. S.C. carried out the modelling of a thin disk to the velocity fields. R.S. led the telescope proposal to obtain the dataset, with a number of key ideas contributed by R.J.B. R.S. was responsible for making all of the figures and writing most of the text. R.J.B., S.C., P.A.O., I.L., G.D.I., P.v.d.W., L.D.B., V.G., J.A.H., B.W.H., R.M. and W.Z. contributed feedback on the various versions of this manuscript and on the telescope proposal.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to R.S. (rs940@cam.ac.uk).

METHODS

Definitions. Throughout this paper we adopt a Chabrier³¹ initial mass function (IMF). For ease of comparison with previous studies, we take H_0 (the Hubble constant) to be $70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, Ω_m (the matter density) to be 0.3, and Ω_Λ (the dark-energy density) to be 0.7, which gives a physical scale of 5.3 kpc per pixel at $z = 6.8$. Magnitudes are quoted in the AB system³². Units are given in terms of solar mass (where $M_\odot = 1.99 \times 10^{33} \text{ g}$) and solar luminosity (where $L_\odot = 3.84 \times 10^{33} \text{ erg s}^{-1}$) where possible.

Data. We obtained ALMA observations centred on the sources COS-3018555981 (right ascension (RA) = 10h 30m 18.5s; declination (dec.) = +02° 15' 59.81") and COS-2987030247 (RA = 10h 0m 29.87s; dec. = +02° 13' 02.47") as part of a filler programme (project code 2015.1.01111.S; principal investigator R.S.) on 14 April 2016, in cycle 3. We requested three tunings to cover the frequency range 1,870.74–1,971.43 GHz in band 6, in order to scan for [C II] at redshift $z = 6.45\text{--}6.90$, corresponding to the 99% photometric redshift probability range⁷. One tuning was executed, scanning the redshift range $z_{\text{[C II]}} = 6.74\text{--}6.90$, with 24 min of source-integration time for each of the targets. The precipitable water vapour (PWV) of the observations was 1.34 mm. The array consisted of 36 antennas and three spectral windows having a bandwidth of 1.875 GHz, to cover a frequency range of 4.95 GHz in a single sideband.

We calibrated and reduced the data with Common Astronomy Software Application (CASA)³³ version 4.5.3, using the automated pipeline, and we imaged the data with the CLEAN task (requiring no iterations, as no continuum sources are detected in the data), using a natural weighting for optimal signal-to-noise. The resulting observations reached an image root-mean-square (r.m.s.) sensitivity of $0.32 \text{ mJy beam}^{-1}$ at 243 GHz in a 50 km s^{-1} channel in both pointings. The primary beam has a resolution of $1.1'' \times 0.7''$ (position angle -48°) for both targets.

We also made use of Hubble Space Telescope (HST) WFC3/F160W (H_{160}) imaging, as well as the photometry of these objects that was used in the selection of our galaxies previously⁷.

Line detections. *COS-3018555981.* We extracted a spectrum from the ALMA cube that was centred on the rest-frame UV continuum of the galaxy detected in the HST H_{160} band of COS-3018555981 as a first guess, and found a clear line detected at around 242 GHz, removed from any atmospheric absorption features and with a peak flux of more than 3.5σ above the local noise. Next, we extracted a spectrally averaged map between 241.85 GHz and 242.10 GHz; this map revealed that the emission line was centred on a faint wing of the UV-continuum detection, $0.27''$ removed from the brightest UV clump (Fig. 1). This offset is similar to the typical uncertainty in the HST astrometry of $0.2''$ (ref. 34); however, if instead the offset is real, this could quite reasonably suggest that the brightest star-forming region in the UV does not spatially coincide with the dynamic centre of the system.

We determined the significance of the detection by measuring the flux on the spectral-line-averaged map in a $1.1'' \times 0.7''$ aperture corresponding to the full-width at half-maximum (FWHM) of the beam, and we repeated this measurement 9,000 times at randomly selected positions of the image, resulting in an estimated signal-to-noise ratio of 8.2. To determine the redshift of COS-3018555981, we extracted a new one-dimensional spectrum from all pixels above the half-maximum of the line detection on the spectral-line-averaged map, and we fitted a Gaussian to the observed line to determine a line centre of $241.97 \pm 0.01 \text{ GHz}$, corresponding to a [C II] redshift of 6.8540 ± 0.0003 , and a linewidth of $232 \pm 30 \text{ km}^{-1}$ FWHM (Fig. 1). The only lines other than [C II] $\lambda = 158 \mu\text{m}$ that are expected to be bright enough to be able to explain our detection are [O I] $\lambda = 63 \mu\text{m}$ and [O III] $\lambda = 88 \mu\text{m}$. However, the [O I] $\lambda = 63 \mu\text{m}$ and [O III] $\lambda = 88 \mu\text{m}$ redshifts of 18.6 and 13.02, respectively, are inconsistent with the HST photometry for this source⁷. Furthermore, the photometric redshift of 6.76 ± 0.07 (ref. 7) is also inconsistent with the [O I] $\lambda = 145 \mu\text{m}$ redshift of 7.5, which is the closest infrared line in frequency, if many times fainter, to [C II] $\lambda = 158 \mu\text{m}$.

COS-2987030247. Similarly to the procedure for COS-3018555981, we first searched for an emission line in the spectrum extracted over the rest-frame UV continuum of COS-2987030247. We found a tentative narrow line at 243.4 GHz—40 MHz removed from an atmospheric absorption feature at 243.5 GHz, where the r.m.s. is 1.5 times greater than the median r.m.s. in the data cube. The spectral-line-averaged map extracted between 243.35 GHz and 243.45 GHz shows a $>5\sigma$ detection close to the position of the HST counterpart; that is, the peak of the map is $0.17''$ removed from the UV-continuum emission (Fig. 1).

By sampling the noise in the spectral-line-averaged map in ellipsoidal apertures of the beam size, we measured a signal-to-noise ratio of 5.1 for the detected line at 243.5 GHz, suggesting that the line is indeed a real detection. To further test the significance of the line we performed a blind line search of the data cube. For each pixel in the cube we extracted a one-dimensional spectrum from averaging all pixels within the ellipsoidal aperture of the beam size, and we fitted any

tentative lines in the spectrum with a Gaussian. If the difference between the χ^2 of the line fit and that of a straight line was greater than 25 (that is, 5σ), we extracted a velocity-averaged image over the FWHM of the line and inspected the significance of the detection on this image. To remove spurious line detections, we again assessed the significance of any potential line from the random sampling of the flux in ellipsoidal apertures on the line map. While we robustly detected the line over COS-2987030247, we found no other sources with a $>5\sigma$ detection in both the one-dimensional spectrum and the spectral-line-averaged map. This test, in combination with the small spatial offset from our HST target, confirms that our line detection over COS-2987030247 is real, and not due to a spurious detection showing up close to the r.m.s. peak of the atmospheric absorption feature.

We extracted a new spectrum from all pixels with a flux above the half-maximum flux in the spectral-line-averaged map, and used this to measure a spectroscopic redshift of $z_{\text{[C II]}} = 6.8076 \pm 0.0002$ for COS-2987030247, in good agreement with the photometric redshift of $z_{\text{phot}} = 6.66 \pm 0.14$.

Dust. We obtained dust continuum measurements after identifying the [C II] line in our data, by averaging the remaining part of the data cubes in frequency. We did not find any evidence for flux above the 1σ noise level in the mean continuum image at the source positions. Therefore, we put an upper limit on the continuum flux, and assumed a grey-body approximation for the dust continuum by considering a range of infrared slopes where we varied both the slope (in the range $\beta_{\text{IR}} = 1\text{--}2$) and the dust temperature (in the range $T_{\text{dust}} = 20\text{--}60 \text{ K}$). We derived a 3σ upper limit on the infrared luminosity of $1.3 \times 10^{11} L_\odot$ and $1.1 \times 10^{11} L_\odot$ for COS-3018555981 and COS-2987030247, respectively.

Given that the UV continuum of galaxies is substantially attenuated by even small amounts of dust, comparing the UV colour and the infrared excess— $\text{IRX} = L_{\text{UV}}/L_{\text{IR}}$ —can provide insights into the dust-attenuation curve in these galaxies. We derived the UV-continuum slope β_{UV} , where the flux density is $f_\lambda \propto \lambda^{\beta_{\text{UV}}}$, from a power-law fit to the HST J_{125} and H_{160} photometry; we found values of -1.22 ± 0.51 and -1.18 ± 0.53 for COS-3018555981 and COS-2987030247, respectively. Often, interpreting the infrared excess as a means to constrain the dust-attenuation curve can be affected by the geometry of the dust³⁵. In particular, a spatial offset between dust-obscured star-forming regions and unobscured UV-emitting regions can produce bluer UV colours for a given IRX ³⁶. The small spatial offsets measured between the UV continuum and [C II] emission in our sources might indicate such an effect of dust geometry here. However, given that our sources already appear much redder than would be predicted by the Meurer¹³ relation for a given IRX , our conclusions are not affected by any spatial offsets of the dust continuum with respect to the UV light.

Star-formation rate and stellar mass. We obtained constraints on the UV-based SFRs rates from the J_{125} band photometry (corresponding to the rest-frame at around $1,600 \text{ \AA}$), and on the infrared-based SFRs from the upper limits on the infrared luminosity, and we converted from luminosity to SFRs using the Kennicutt³⁷ scaling relations. For COS-3018555981, a foreground object of $z = 0.74$ is visible at a projected distance of $2.6''$, which could introduce a small boost to the measured fluxes owing to gravitational lensing. However, the stellar mass of this object is only $4 \times 10^9 M_\odot$ (ref. 38), which suggests a modest halo mass, and therefore we estimate the magnification of this source to be no more than 0.1 magnitude (that is, no larger than the measured random errors), as discussed recently²².

Using the deconvolved size of the [C II] emission as the size of the galaxy, we found a SFR density, Σ_{SFR} , of $0.91 M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$ and $0.75 M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$ for COS-3018555981 and COS-2987030247, respectively. This is in good agreement with the SFRs obtained using [C II] as a spatially resolved SFR indicator, using the relation calibrated for galaxies from the local KINGFISH sample²⁰, which predicts a Σ_{SFR} of $0.68 M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$ and $0.34 M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$ on the basis of the [C II] surface brightness, $\Sigma_{\text{[C II]}}$, of $8.5 \times 10^{40} \text{ erg s}^{-1} \text{ kpc}^{-2}$ and $4.6 \times 10^{40} \text{ erg s}^{-1} \text{ kpc}^{-2}$.

Although the rest-frame optical photometry of $z > 4$ galaxies can be heavily affected by strong nebular emission lines^{39,40}, the redshift range $z \approx 6.6\text{--}7.0$ offers a unique window where the 4.5 Spitzer/IRAC band is free from contamination by nebular emission lines^{7,8}, providing a good constraint with which to model the stellar population of galaxies at these redshifts. We used the Bayesian code MAGPHYS⁴¹ with the HIGHZ extension⁴² to fit the stellar population. We included the continuum constraints at 243 GHz, but we removed the 3.6- μm Spitzer/IRAC photometry, as this band is affected by high equivalent-width nebular emission ($\text{EW}_{\text{[O III]} + \text{H}\beta}$ is about $1,000\text{--}1,500 \text{ \AA}$; ref. 7). We find that both galaxies have best-fitting stellar masses of about $(1\text{--}2) \times 10^9 M_\odot$.

Velocity structure and dynamic mass. The line maps extracted in Fig. 1 suggest that the [C II] emission is spatially resolved in both galaxies, which allows us to investigate the presence of any velocity structure in these galaxies. For the central $4''$ of the data cube, we extracted a one-dimensional spectrum at every pixel, by averaging all the flux within an elliptical aperture the size of the beam centred on the pixel. We fitted a Gaussian to these spectra, using the parameters from the fit

to the integrated spectrum as initial parameters. We required the fit to the one-dimensional spectrum to be significant at $>5\sigma$.

We measured a projected velocity difference over the galaxies of $111 \pm 28 \text{ km}^{-1}$ and $54 \pm 20 \text{ km}^{-1}$ for COS-3018555981 and COS-2987030247, respectively, using the minimum and maximum central frequencies taken from the fits that are significant at $>5\sigma$. Galaxies with $\Delta v_{\text{obs}}/2\sigma_{\text{tot}}$ ratios greater than 0.4 (using the measured line widths in Extended Data Table 1 to estimate the integrated velocity dispersion) can be classified as probable rotation-dominated systems in cases where the data quality prevents reliable kinematic modelling²⁵. This is an approximate diagnostic based on simulations of disk galaxies with a wide range of intrinsic properties. The observed limit of $\Delta v_{\text{obs}}/2\sigma_{\text{tot}}$ at around 0.4 corresponds to the intrinsic ratio of $v_{\text{rot}}/\sigma_0 = 1$ (ref. 25). We tested the robustness of the observed velocity gradient by re-imaging the ALMA data with CASA, using a Briggs weighting with a robustness parameter of 0.5, which produces images of the [C II] emission at a lower signal-to-noise ratio but slightly improved spatial resolution ($0.9'' \times 0.7''$). We confirmed that the same analysis on the higher-resolution data still produced a velocity gradient with the same projected velocity difference over the two galaxies.

We assumed that these galaxies can be described as symmetric rotating disks. This is a reasonable assumption given the consistent prediction of high-resolution hydrodynamic zoom simulations, which show that cool gas indeed settles into regular rotating disks^{27,43–45}, and given the prevalence of disks among star-forming galaxies at lower redshifts^{46–48}. To derive a dynamic mass for these systems, we adopted two methods. First, we use the approximation that the dynamic mass is estimated from $M_{\text{dyn}}(r < r_{1/2}) = (v_d^2 r_{1/2})/G$, where $r_{1/2}$ is the half-light radius of [C II], G is the gravitational constant, and v_d is derived from the average of the observed velocity gradient over the galaxy, $v_d \sin(i) = 1.3 \Delta v_{\text{obs}}$ (where i is the disk inclination), and the integrated velocity dispersion, $v_d \sin(i) = 0.99 \sigma_{\text{tot}}$ (ref. 25). We estimated a half-light radius ($r_{1/2}$) and the inclination of the system ($\sin(i)$) from an ellipsoidal fit to the [C II] emission line map using CASA (corrected for the beam), and found $r_{1/2}$ values of $2.6 \pm 0.8 \text{ kpc}$ and $3.1 \pm 1.0 \text{ kpc}$, and $\sin(i)$ values of 0.59 ± 0.15 and 0.88 ± 0.06 , for our sources. We derived dynamic masses of $(25.3 \pm 15.4) \times 10^9 M_{\odot}$ and $(3.4 \pm 1.7) \times 10^9 M_{\odot}$ for COS-3018555981 and COS-2987030247, respectively.

To obtain a second mass estimate, we modelled the velocity field by assuming that the gas is rotating in a circularly symmetric thin disk, with a gravitational potential that depends only on the disk mass and assuming an exponential distribution of the surface mass density. The circular velocity is projected along the line of sight, weighted by the profile of the intrinsic line surface brightness, and convolved with the beam size of the observations. Free parameters of our model are the inclination of the disk, the position angle of the disk line of nodes, the systemic velocity of the galaxy, and the dynamic mass, measured in a radius of 5 kpc. Our method has been successfully applied to ALMA observations of [C II] emitting sources at redshifts of around 5 (refs 49, 50). Our free parameters were simultaneously constrained from the velocity maps using least-squares fitting. Furthermore, we fitted the coordinates of the disk centre on the basis of the surface brightness maps, which are a minor uncertainty to our final results. We estimated uncertainties from the χ^2 parameter space, which was constrained with Monte Carlo Markov chain simulations. The best-fitting model describes our velocity field well, leaving small residuals (Extended Data Fig. 1). The best-fit parameters indicate half-light radii of $1.7^{+0.4}_{-0.3} \text{ kpc}$ and $2.1^{+2.1}_{-1.1} \text{ kpc}$, inclination angles of $\sin(i) = 0.87^{+0.07}_{-0.10}$ and $0.64^{+0.22}_{-0.30}$, and dynamic masses of $1.0^{+0.3}_{-0.2} \times 10^{10} M_{\odot}$ and $0.4^{+0.9}_{-0.3} \times 10^{10} M_{\odot}$ for COS-3018555981 and COS-2987030247, respectively. These values are all consistent (within the uncertainties) with the estimates derived above. We therefore adopted this more sophisticated method for our fiducial dynamic mass estimates.

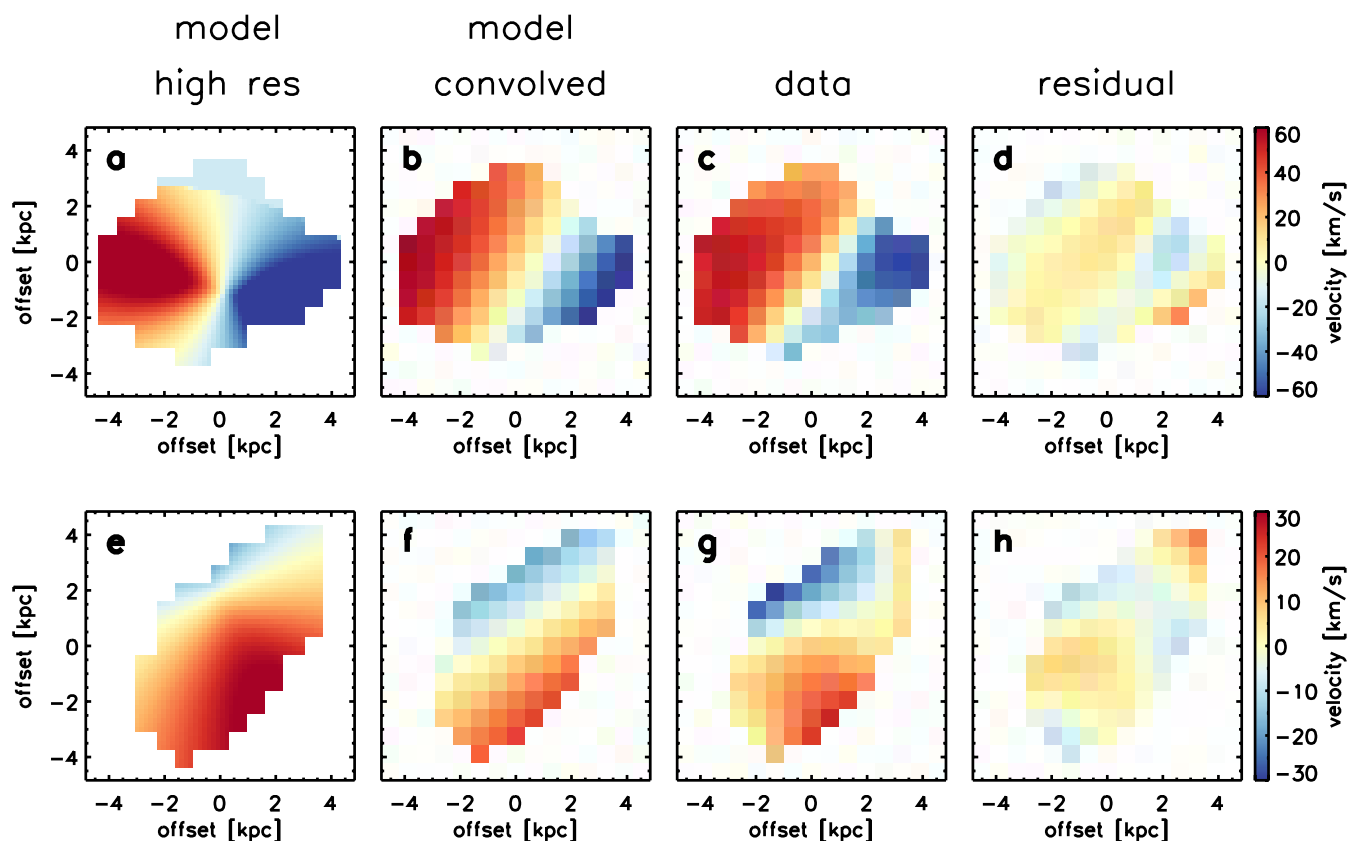
In the methods described above, the effect of turbulence on the estimated dynamic masses is not included^{51,52}. For dispersion-dominated galaxies, the dynamic mass (including pressure support) can be estimated by $M_{\text{dyn}} = 2R_{1/2}(v_{\text{rot}}^2 + \sigma_0^2)/G$ (ref. 53), where v_{rot} is the inclination-corrected velocity gradient, and we estimate σ_0 values of 55 km^{-1} and 30 km^{-1} . The resulting dynamical masses are 0.3 dex and 0.4 dex higher than our previous estimates for COS-3018555981 and COS-2987030247, respectively. To study the effect of

asymmetric drift on the rotation curve in more detail, higher-resolution observations will be required.

Code availability. The data used here were reduced and partly analysed with the public code CASA, available at https://casa.nrao.edu/casa_obtaining.shtml. The reduction pipeline for this source can be downloaded as part of the ALMA observations with project code 2015.1.01111.S, available in the archive at <https://almascience.nrao.edu/alma-data/archive>. The kinematic models used for this study are available from the corresponding author upon request.

Data availability. The data used in this publication are publicly available in the data archive <https://almascience.nrao.edu/alma-data/archive>, and can be retrieved with the project code 2015.1.01111.S or using the name of the principal investigator, 'Smit, Renske'.

- Chabrier, G. Galactic stellar and substellar initial mass function. *Publ. Astron. Soc. Pac.* **115**, 763–795 (2003).
- Oke, J. B. & Gunn, J. E. Secondary standard stars for absolute spectrophotometry. *Astrophys. J.* **266**, 713–717 (1983).
- McMullin, J. P., Waters, B., Schiebel, D., Young, W. & Golap, K. CASA architecture and applications. In *Astronomical Data Analysis Software and Systems XVI* (eds Shaw, R. A., Hill, F. & Bell, D. J.) 127–130 (*Astron. Soc. Pac. Conf. Ser.* Vol. 376, 2007).
- Dunlop, J. S. *et al.* A deep ALMA image of the Hubble ultra deep field. *Mon. Not. R. Astron. Soc.* **466**, 861–883 (2017).
- Narayanan, D. *et al.* The IRX-beta dust attenuation relation in cosmological galaxy formation simulations. Preprint at <https://arxiv.org/abs/1705.05858> (2017).
- Koprowski, M. P. *et al.* A resolved map of the infrared excess in a Lyman break galaxy at $z = 3$. *Astrophys. J.* **828**, L21 (2016).
- Kennicutt, R. C. & Evans, N. J. Star formation in the Milky Way and nearby galaxies. *Annu. Rev. Astron. Astrophys.* **50**, 531–608 (2012).
- Brammer, G. B. *et al.* 3D-HST: a wide-field grism spectroscopic survey with the Hubble space telescope. *Astrophys. J. Suppl. Ser.* **200**, 13 (2012).
- Stark, D. P. *et al.* Keck spectroscopy of $3 < z < 7$ faint Lyman break galaxies: the importance of nebular emission in understanding the specific star formation rate and stellar mass density. *Astrophys. J.* **763**, 129 (2013).
- Smit, R. *et al.* Inferred $\text{H}\alpha$ flux as a star formation rate indicator at $z \sim 4$ –5: implications for dust properties, burstiness, and the $z = 4$ –8 star formation rate functions. *Astrophys. J.* **833**, 254 (2016).
- da Cunha, E., Charlot, S. & Elbaz, D. A simple model to interpret the ultraviolet, optical and infrared emission from galaxies. *Mon. Not. R. Astron. Soc.* **388**, 1595–1617 (2008).
- da Cunha, E. *et al.* An ALMA survey of sub-millimeter galaxies in the extended Chandra deep field south: physical properties derived from ultraviolet-to-radio modeling. *Astrophys. J.* **806**, 110 (2015).
- Feng, Y. *et al.* The formation of Milky Way-mass disk galaxies in the first 500 million years of a cold dark matter universe. *Astrophys. J.* **808**, L17 (2015).
- Katz, H., Kimm, T., Sijacki, D. & Haehnelt, M. G. Interpreting ALMA observations of the ISM during the epoch of reionization. *Mon. Not. R. Astron. Soc.* **468**, 4831–4861 (2017).
- Pallottini, A. *et al.* Zooming on the internal structure of $z \sim 6$ galaxies. *Mon. Not. R. Astron. Soc.* **465**, 2540–2558 (2017).
- Stott, J. P. *et al.* The KMOS redshift one spectroscopic survey (KROSS): dynamical properties, gas and dark matter fractions of typical $z \sim 1$ star-forming galaxies. *Mon. Not. R. Astron. Soc.* **457**, 1888–1904 (2016).
- Wuyts, S. *et al.* KMOS3D: dynamical constraints on the mass budget in early star-forming disks. *Astrophys. J.* **831**, 149 (2016).
- Mason, C. A. *et al.* First results from the KMOS lens-amplified spectroscopic survey (KLASS): kinematics of lensed galaxies at cosmic noon. *Astrophys. J.* **838**, 14 (2017).
- Carniani, S. *et al.* Strongly star-forming rotating disks in a complex merging system at $z = 4.7$ as revealed by ALMA. *Astron. Astrophys.* **559**, A29 (2013).
- Williams, R. J. *et al.* Constraining the nature of two $\text{Ly}\alpha$ emitters detected by ALMA at $z = 4.7$. *Mon. Not. R. Astron. Soc.* **439**, 2096–2101 (2014).
- Burkert, A. *et al.* High-redshift star-forming galaxies: angular momentum and baryon fraction, turbulent pressure effects, and the origin of turbulence. *Astrophys. J.* **725**, 2324–2332 (2010).
- Burkert, A. *et al.* The angular momentum distribution and baryon content of star-forming galaxies at $z \sim 1$ –3. *Astrophys. J.* **826**, 214 (2016).
- Newman, S. F. *et al.* The SINS/zC-SINF survey of $z \sim 2$ galaxy kinematics: the nature of dispersion-dominated galaxies. *Astrophys. J.* **767**, 104 (2013).



Extended Data Figure 1 | Models of the velocity fields of COS-3018555981 and COS-2987030247, using a disk model.
a–h, Model fits to the velocity gradients in COS-3018555981 (**a–d**) and COS-2987030247 (**e–h**), assuming that the gas is rotating in an exponential, circularly symmetric thin disk. **a**, **e**, High-resolution disk

model before convolution with the beam; **b**, **f**, disk model at the resolution of our observations; **c**, **g**, our velocity maps, as shown in Fig. 3; **d**, **h**, residuals after subtraction of the model. Although the disk model is not a unique solution for these velocity fields, our galaxies are well described by regular rotation.

Extended Data Table 1 | Galaxy properties

ID	COS-3018555981	COS-2987030247
z_{phot}^7	6.76±0.07	6.66±0.14
$z_{\text{[CII]}}^*$	6.8540±0.0003	6.8076±0.0002
S/N [†]	8.2	5.1
[C II] line flux (Jy km s ⁻¹)*	0.39±0.05	0.31±0.04
FWHM _[CII] (km s ⁻¹)*	232±30	124±18
158μm continuum flux (μJy)	<87 [‡]	<75 [‡]
$L_{\text{[CII]}}$ (10 ⁸ L _⊙)	4.7±0.5	3.6±0.5
L_{UV} (10 ¹¹ L _⊙)	1.1±0.1	1.3±0.1
L_{IR} (10 ¹¹ L _⊙)	<1.3 [‡]	<1.1 [‡]
SFR _{IR} (M _⊙ yr ⁻¹)	<19 [‡]	<16 [‡]
SFR _{UV} (M _⊙ yr ⁻¹)	19.2±1.6	22.7±2.0
M_* (10 ⁹ M _⊙)	1.4 ^{+0.7} _{-0.2}	1.7 ^{+0.5} _{-0.2}
M_{dyn} (10 ⁹ M _⊙)	10 ⁺³ ₋₂	4 ⁺⁹ ₋₃
$\Delta v_{\text{obs}} / 2\sigma_{\text{tot}}$	0.57±0.16	0.52±0.21
$r_{1/2, \text{[CII]}}$ (kpc)	2.6±0.8	3.1±1.0
β_{UV}	-1.22±0.51	-1.18±0.53
EW([O III]+Hβ) (Å) ⁷	1424±143	1128±166

*These values were measured from a Gaussian fit to the integrated spectrum within the half-peak-power contour.

[†]The signal-to-noise ratio (S/N) was measured in a beam-sized aperture (centred on the HST counterpart) on a velocity-averaged image extracted over the detected line.

[‡]3σ limit.

An extreme magneto-ionic environment associated with the fast radio burst source FRB 121102

D. Michilli^{1,2*}, A. Seymour^{3*}, J. W. T. Hessels^{1,2*}, L. G. Spitler⁴, V. Gajjar^{5,6,7}, A. M. Archibald^{1,2}, G. C. Bower⁸, S. Chatterjee⁹, J. M. Cordes⁹, K. Gourdji², G. H. Heald¹⁰, V. M. Kaspi¹¹, C. J. Law¹², C. Sobey^{10,13}, E. A. K. Adams^{1,14}, C. G. Bassa¹, S. Bogdanov¹⁵, C. Brinkman¹⁶, P. Demorest¹⁷, F. Fernandez³, G. Hellbourg¹², T. J. W. Lazio¹⁸, R. S. Lynch^{19,20}, N. Maddox¹, B. Marcote²¹, M. A. McLaughlin^{20,22}, Z. Paragi²¹, S. M. Ransom²³, P. Scholz²⁴, A. P. V. Siemion^{12,25,26}, S. P. Tendulkar¹¹, P. Van Rooy²⁷, R. S. Wharton⁴ & D. Whitlow³

Fast radio bursts are millisecond-duration, extragalactic radio flashes of unknown physical origin^{1–3}. The only known repeating fast radio burst source^{4–6}—FRB 121102—has been localized to a star-forming region in a dwarf galaxy^{7–9} at redshift 0.193 and is spatially coincident with a compact, persistent radio source^{7,10}. The origin of the bursts, the nature of the persistent source and the properties of the local environment are still unclear. Here we report observations of FRB 121102 that show almost 100 per cent linearly polarized emission at a very high and variable Faraday rotation measure in the source frame (varying from $+1.46 \times 10^5$ radians per square metre to $+1.33 \times 10^5$ radians per square metre at epochs separated by seven months) and narrow (below 30 microseconds) temporal structure. The large and variable rotation measure demonstrates that FRB 121102 is in an extreme and dynamic magneto-ionic environment, and the short durations of the bursts suggest a neutron star origin. Such large rotation measures have hitherto been observed^{11,12} only in the vicinities of massive black holes (larger than about 10,000 solar masses). Indeed, the properties of the persistent radio source are compatible with those of a low-luminosity, accreting massive black hole¹⁰. The bursts may therefore come from a neutron star in such an environment or could be explained by other models, such as a highly magnetized wind nebula¹³ or supernova remnant¹⁴ surrounding a young neutron star.

Using the 305-m William E. Gordon Telescope at the Arecibo Observatory, we detected 16 bursts from FRB 121102 at radio frequencies in the range 4.1–4.9 GHz (Table 1). Complete polarization parameters were recorded at a 10.24- μ s time resolution. See Methods and Extended Data Figs 1–6 for observation and analysis details.

The 4.5-GHz bursts have typical widths smaller than about 1 ms, which are narrower than the 2–9-ms bursts previously detected at lower frequencies^{5,15}. In some cases they show multiple components and structures close to the sampling time of the data. Burst 6 (Table 1) is particularly striking, with a width smaller than about 30 μ s (which constrains the size of the emitting region to below about 10 km, assuming no other geometric or relativistic effects). The evolution of burst

morphology with frequency complicates the determination⁵ of the dispersion measure ($DM = \int_0^d n_e(l)dl$, where d is the distance to the source in parsec, l is the line-of-sight position and n_e is the electron density in electrons per cubic centimetre), but aligning the narrow component in burst 6 results in $DM = 559.7 \pm 0.1 \text{ pc cm}^{-3}$, which is consistent^{4,5,15,16} with other bursts detected since 2012 and suggests that any real dispersion measure variations are below the level of about 1%.

After correcting for Faraday rotation and accounting for about 2% depolarization from the finite channel widths, the bursts are consistently linearly polarized to about 100% (Fig. 1). The polarization angles $PA = PA_\infty + \theta$ (where PA_∞ is a reference angle at infinite frequency, $\theta = RM\lambda^2$ is the rotation angle of the electric field vector, RM is the Faraday rotation measure and λ is the observing wavelength) are flat across the observed frequency range and burst envelopes (ΔPA smaller than about 5° ms^{-1}). This could mean that the burst durations reflect the timescale of the emission process and not the rate of a rotating beam sweeping across the line of sight. Any circular polarization is lower than a few per cent of the total intensity. The Faraday rotation measure is defined as $RM = 0.81 \int_d B_\parallel(l) n_e(l) dl$, where B_\parallel is the line-of-sight magnetic field strength (in microgauss); by convention, the rotation measure is positive when the magnetic field points towards the observer. On average, the observed rotation measure is $RM_{\text{obs}} = (+1.027 \pm 0.001) \times 10^5 \text{ rad m}^{-2}$ and varies by about 0.5% between Arecibo observing sessions spanning a month (Fig. 2; Table 1). The lack of polarization in previous burst detections^{15,16} at 1.1–2.4 GHz is consistent with the relatively coarse frequency channels that cause bandwidth depolarization and constrains $|RM_{\text{obs}}|$ to above about 10^4 rad m^{-2} at those epochs.

Confirmation of this extreme Faraday rotation comes from independent observations at 4–8 GHz with the 110-m Robert C. Byrd Green Bank Telescope (GBT), which give $RM_{\text{obs}} = (+0.935 \pm 0.001) \times 10^5 \text{ rad m}^{-2}$ at an epoch seven months after the Arecibo detections. The GBT and Arecibo RM_{obs} values differ with high

¹ASTRON, Netherlands Institute for Radio Astronomy, Postbus 2, 7990 AA Dwingeloo, The Netherlands. ²Anton Pannekoek Institute for Astronomy, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands. ³National Astronomy and Ionosphere Center, Arecibo Observatory, Puerto Rico 00612, USA. ⁴Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany. ⁵Space Science Laboratory, 7 Gauss Way, University of California, Berkeley, California 94710, USA. ⁶Xinjiang Astronomical Observatory, CAS, 150 Science 1-Street, Urumqi, Xinjiang 830011, China. ⁷Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Nanjing 210008, China. ⁸Academia Sinica Institute of Astronomy and Astrophysics, 645 N. A'ohoku Place, Hilo, Hawaii 96720, USA. ⁹Cornell Center for Astrophysics and Planetary Science and Department of Astronomy, Cornell University, Ithaca, New York 14853, USA. ¹⁰CSIRO Astronomy and Space Science, 26 Dick Perry Avenue, Kensington, Western Australia 6151, Australia. ¹¹Department of Physics and McGill Space Institute, McGill University, 3600 University, Montréal, Quebec H3A 2T8, Canada. ¹²Department of Astronomy and Radio Astronomy Lab, University of California, Berkeley, California 94720, USA. ¹³International Centre for Radio Astronomy Research - Curtin University, GPO Box U1987, Perth, Western Australia 6845, Australia. ¹⁴Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AA Groningen, The Netherlands. ¹⁵Columbia Astrophysics Laboratory, Columbia University, New York, New York 10027, USA. ¹⁶Physics Department, University of Vermont, Burlington, Vermont 05401, USA. ¹⁷National Radio Astronomy Observatory, PO Box O, Socorro, New Mexico 87801 USA. ¹⁸Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA. ¹⁹Green Bank Observatory, PO Box 2, Green Bank, West Virginia 24944, USA. ²⁰Center for Gravitational Waves and Cosmology, Chestnut Ridge Research Building, Morgantown, West Virginia 26505, USA. ²¹Joint Institute for VLBI ERIC, Postbus 2, 7990 AA Dwingeloo, The Netherlands. ²²Department of Physics and Astronomy, West Virginia University, Morgantown, West Virginia 26506, USA. ²³National Radio Astronomy Observatory, Charlottesville, Virginia 22903, USA. ²⁴National Research Council of Canada, Herzberg Astronomy and Astrophysics, Dominion Radio Astrophysical Observatory, PO Box 248, Penticton, British Columbia V2A 6J9, Canada. ²⁵Radboud University, Nijmegen, Comeniuslaan 4, 6525 HP Nijmegen, The Netherlands. ²⁶SETI Institute, 189 North Bernardo Avenue 200, Mountain View, California 94043, USA. ²⁷Department of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, Ohio 44106, USA.

*These authors contributed equally to this work.

Table 1 | Properties of Arecibo (1–16) and GBT (GBT-1 and GBT-2) bursts

Burst	Modified Julian date	Width (ms)	S (Jy)	F (Jy ms)	RM_{obs} (rad m^{-2})	PA_{∞} ($^{\circ}$)	RM_{global} (rad m^{-2})	$PA_{\infty}^{\text{global}}$ ($^{\circ}$)
1	57,747.1295649013	0.80	0.9	0.7	$+102,741 \pm 9$	49 ± 2		
2	57,747.1371866766	0.85	0.3	0.2	$+102,732 \pm 34$	55 ± 9		
3	57,747.1462710273	0.22	0.8	0.2	$+102,689 \pm 18$	64 ± 5		
4	57,747.1515739398	0.55	0.2	0.09	–	–		
5	57,747.1544674919	0.76	0.2	0.1	–	–	$+102,708 \pm 4$	
6	57,747.1602892954	0.03	1.8	0.05	$+102,739 \pm 35$	49 ± 9		
7	57,747.1603436945	0.31	0.6	0.2	$+102,663 \pm 33$	71 ± 9		
8	57,747.1658277033	1.36	0.4	0.5	$+102,668 \pm 18$	67 ± 4		
9	57,747.1663749941	1.92	0.2	0.3	–	–		58 ± 1
10	57,747.1759674338	0.98	0.2	0.2	–	–		
11	57,748.1256436428	0.95	0.1	0.1	–	–		
12	57,748.1535244366	0.42	0.4	0.2	$+102,508 \pm 35$	63 ± 10		
13	57,748.1552149312	0.78	0.8	0.6	$+102,522 \pm 17$	59 ± 4	$+102,521 \pm 4$	
14	57,748.1576076618	0.15	1.2	0.2	$+102,489 \pm 18$	67 ± 5		
15	57,748.1756968287	0.54	0.4	0.4	$+102,492 \pm 37$	64 ± 10		
16	57,772.1290302972	0.74	0.8	0.6	$+103,020 \pm 12$	64 ± 3	$+103,039 \pm 4$	
GBT-1	57,991.5801286366	0.59	0.4	0.2	$+93,526 \pm 72$	73 ± 8	$+93,573 \pm 24$	68 ± 2
GBT-2	57,991.5833032369	0.27	0.9	0.2	$+93,533 \pm 42$	71 ± 4		

Modified Julian dates are referenced to infinite frequency at the Solar System barycentre; their uncertainties are of the order of the burst widths. Widths have uncertainties of about $10 \mu\text{s}$. Peak flux densities S and fluences F have about 20% fractional uncertainties. Rotation measures are not corrected for redshift, and polarization angles are referenced to infinite frequency. Bursts with no individual rotation measure entry (–) were too weak to reliably fit on their own. The last two columns refer to a global fit of all bursts. All errors are 1σ ; see Methods for observational details.

statistical significance and indicate that the rotation measure can vary by at least 10% on half-year timescales (Table 1 and Extended Data Fig. 5).

The Faraday rotation must come almost exclusively from within the host galaxy; the expected Milky Way contribution¹⁷ is $-25 \pm 80 \text{ rad m}^{-2}$, while estimated intergalactic medium contributions¹⁸ are lower than about 10^2 rad m^{-2} . In the source reference frame, $RM_{\text{src}} = RM_{\text{obs}}(1+z)^2 = +1.46 \times 10^5 \text{ rad m}^{-2}$ and $+1.33 \times 10^5 \text{ rad m}^{-2}$ for the Arecibo and GBT data, respectively, where z is the redshift. Without a correspondingly large change in the dispersion measure,

the observed variations in rotation measure indicate that the Faraday rotation comes from a compact region with a high magnetic field. Furthermore, that region must be close to FRB 121102 because it is very unlikely that an unrelated small structure with the required high magnetic field is coincidentally in the line of sight.

We can fit all 16 Arecibo bursts with a single polarization angle $PA_{\infty}^{\text{global}} = 58^{\circ} \pm 1^{\circ}$ (referenced to infinite frequency; measured anti-clockwise from North to East) and a single RM_{global} per observation day (Table 1). However, we cannot rule out small changes in the

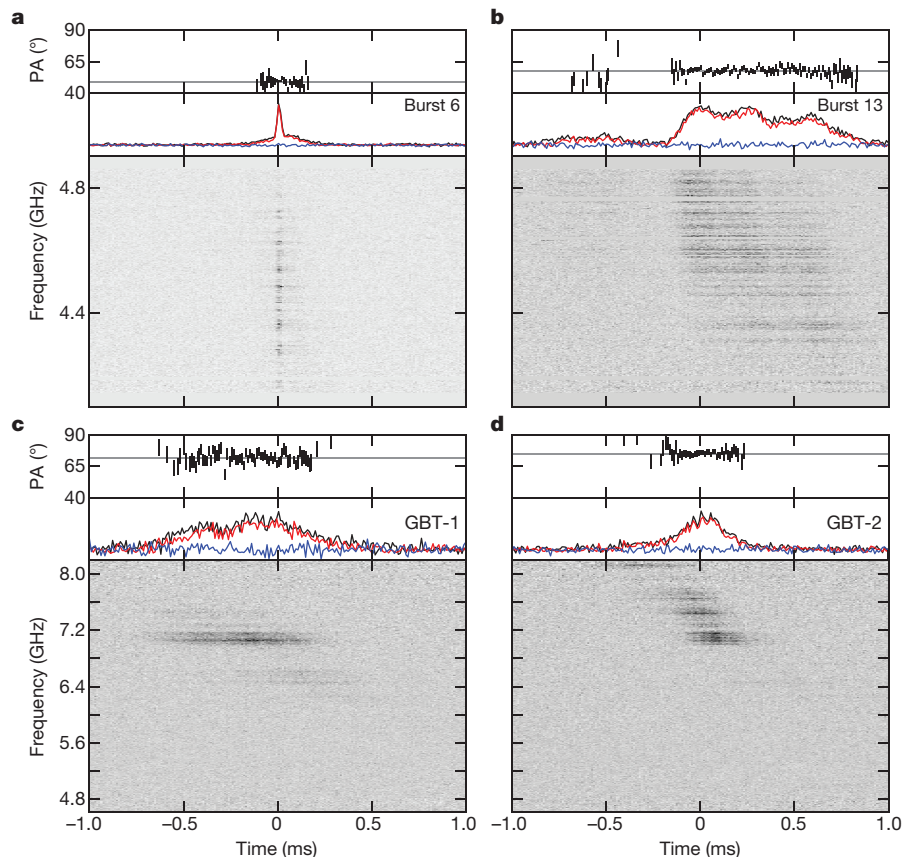


Figure 1 | Polarization angles, pulse profile and spectrum of four bursts. The grey horizontal lines indicate the average polarization angle of each burst. The red and blue lines indicate linear and circular polarization profiles, respectively, while the black line is the total intensity. **a, b,** The

Arecibo bursts are plotted with time and frequency resolutions of $10.24 \mu\text{s}$ and 1.56 MHz , respectively. **c, d,** The GBT bursts are plotted with time and frequency resolutions of $10.24 \mu\text{s}$ and 5.86 MHz , respectively.

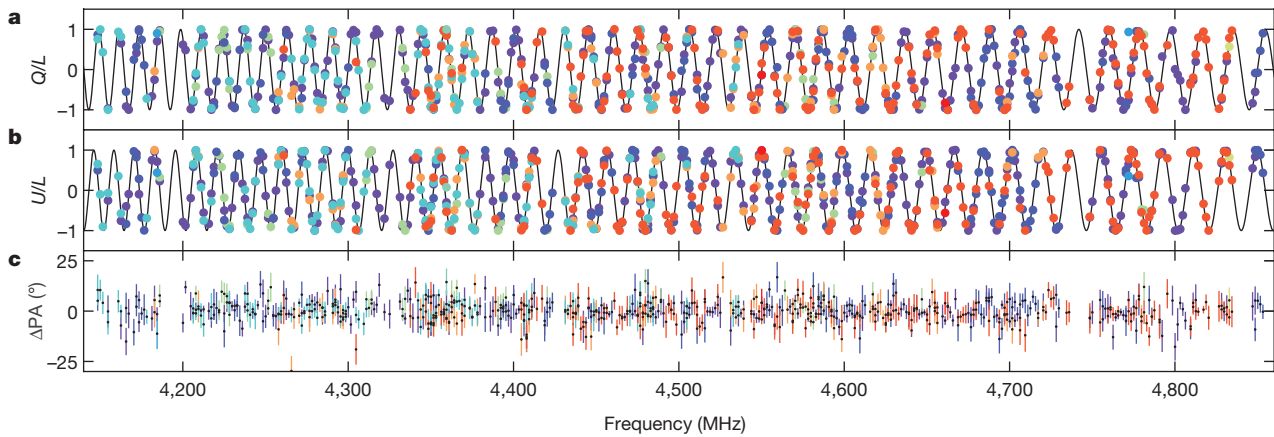


Figure 2 | Faraday rotation in the bursts. **a**, **b**, Variations of the Stokes parameters Q (**a**) and U (**b**) with frequency, normalized by the total linear polarization ($L = \sqrt{Q^2 + U^2}$), for the six brightest Arecibo bursts detected on modified Julian date 57,747. Different bursts are plotted using different colours. Only data points with signal-to-noise ratio higher than 5 are

plotted and do not include uncertainties. The black lines represent the best-fitting Faraday rotation model for the global values reported in Table 1. **c**, Difference between calculated and measured polarization angles (ΔPA) with 1σ uncertainties around the central values, which are indicated with black dots.

rotation measure (below about 50 rad m^{-2}) and polarization angle (lower than approximately 10°) between bursts. The GBT data are not well modelled by the use of a single $PA_{\infty}^{\text{global}}$ value, but this could be an instrumental difference or reflection of the higher observing frequency. The near constancy of the polarization angle suggests that the burst emitter has a stable geometric orientation with respect to the observer. A linear polarization fraction higher than about 98% at a single rotation measure constrains turbulent scatter¹⁹ as $\sigma_{\text{RM}} < 25 \text{ rad m}^{-2}$ and the linear gradient across the source as $\Delta_{\text{RM}} < 20 \text{ rad m}^{-2}$, and there is no evidence of deviations from the squared-wavelength (λ^2) scaling of the Faraday rotation effect. Analysis with the RM Synthesis technique and the deconvolution procedure RMCLEAN also implies a ‘Faraday-thin’ medium (see Methods).

In the rest frame, the host galaxy contributes a dispersion measure $DM_{\text{host}} \approx 70\text{--}270 \text{ pc cm}^{-3}$ to the total dispersion measure of the bursts⁸. Given RM_{src} , this corresponds to an estimated line-of-sight magnetic field $B_{\parallel} = 0.6f_{\text{DM}} - 2.4f_{\text{DM}} \text{ mG}$. This is a lower-limit range because the dispersion measure contribution that is related to the observed rotation measure (DM_{RM}) could be much smaller than the total dispersion measure contribution of the host (DM_{host} , dominated by the star-forming region), which we quantify by the scaling factor $f_{\text{DM}} = DM_{\text{host}}/DM_{\text{RM}} \geq 1$. For comparison, typical magnetic field strengths within the interstellar medium of our Galaxy²⁰ are only about $5 \mu\text{G}$.

We can constrain the electron density, electron temperature (T_e) and length scale (L_{RM}) of the region causing the Faraday rotation by balancing the magnetic field and thermal energy densities (Extended Data Fig. 6). For example, assuming equipartition and $T_e = 10^6 \text{ K}$, we find a density of $n_e \approx 10^2 \text{ cm}^{-3}$ on a length scale of $L_{\text{RM}} \approx 1 \text{ pc}$, comparable to the upper limit of the size of the persistent source¹⁰.

A star-forming region, such as that hosting FRB 121102, will contain H II regions of ionized hydrogen. Although very compact H II regions have sufficiently high magnetic fields and electron densities to explain the large rotation measure, the constraints from DM_{host} and the absence of free-free absorption of the bursts exclude a wide range of H II region sizes and densities²¹ for typical temperatures of 10^4 K .

The environment around a massive black hole is consistent with the n_e , L_{RM} and T_e constraints²², and the properties of the persistent source are compatible with those of a low-luminosity, accreting massive black hole¹⁰. The high rotation measure towards the Galactic Centre magnetar²³ PSR J1745–2900 (Fig. 3), $RM = -7 \times 10^4 \text{ rad m}^{-2}$, provides an intriguing observational analogy for a scenario in which the bursts are produced by a neutron star in the immediate environment of a massive black hole. However, the bursts of FRB 121102 are many orders of magnitude more energetic than those of PSR J1745–2900 or any Galactic pulsar.

An alternative description of FRB 121102 has been proposed by a millisecond magnetar model^{8,10,13}. According to that model, one would expect a surrounding supernova remnant and nebula powered by the central neutron star. The n_e , L_{RM} and T_e constraints are broadly compatible with the conditions in pulsar-wind nebulae, but dense filaments like those seen in the Crab Nebula²⁴ may need to be invoked to explain the high and variable rotation measure of FRB 121102. In a young neutron star scenario, an expanding supernova remnant could also in principle produce a high rotation measure by sweeping up surrounding ambient medium and progenitor ejecta²⁵. A more detailed discussion of these scenarios is provided in Methods, and more exotic models also remain possible²⁶.

Regardless of its nature, FRB 121102 clearly inhabits an extreme magneto-ionic environment. In contrast, Galactic pulsars with comparable dispersion measures have rotation measures that are

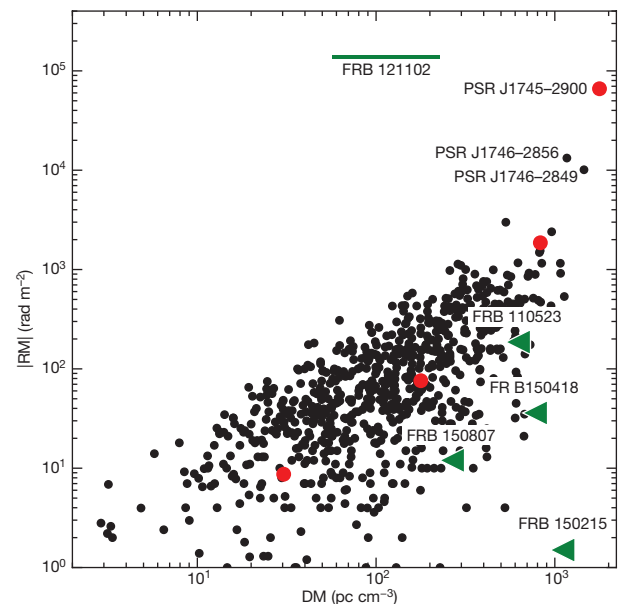


Figure 3 | Magnitude of rotation measure versus dispersion measure for fast radio bursts and Galactic pulsars. Radio-loud magnetars are highlighted with red dots, while radio pulsars and magnetars closest to the Galactic Centre³⁰ are labelled by name. The green bar represents FRB 121102 and the uncertainty on the dispersion measure contribution of the host galaxy⁸. Green triangles are other fast radio bursts with measured rotation measure; here the dispersion measure is the upper limit of the contribution from the host galaxy.

smaller than a hundredth of the RM_{src} value of FRB 121102 (Fig. 3), which is also about 500 times larger than those previously detected in fast radio bursts²⁷. The five other known fast radio bursts with polarimetric measurements present a heterogeneous picture, with a range of polarization fractions and rotation measures³. As previously considered²⁸, the large Faraday rotation of FRB 121102 further suggests that fast radio bursts with no detectable linear polarization may actually have very large $|RM|$, higher than 10^4 – 10^5 rad m⁻², that was undetectable because of the limited frequency resolution (0.4-MHz channels at 1.4 GHz) of the observations.

Monitoring the rotation measure and polarization angle of FRB 121102 with time, along with searches for polarization and Faraday rotation from the persistent source, can help differentiate among competing models. FRB 121102 is unusual not only because of its large rotation measure but also because it is the only known repeating fast radio burst. This may indicate that FRB 121102 is a fundamentally different type of source compared to the rest of the fast radio burst population; future measurements may investigate a possible correlation between fast radio burst repetition and rotation measure. Perhaps the markedly higher activity level of FRB 121102 compared to other known fast radio bursts is predominantly a consequence of its environment; for example, because these magnetized structures can also boost the detectability of the bursts via plasma lensing²⁹.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 November; accepted 20 November 2017.

- Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J. & Crawford, F. A bright millisecond radio burst of extragalactic origin. *Science* **318**, 777–780 (2007).
- Thornton, D. *et al.* A population of fast radio bursts at cosmological distances. *Science* **341**, 53–56 (2013).
- Petroff, E. *et al.* FRBCAT: the fast radio burst catalogue. *Publ. Astron. Soc. Aust.* **33**, e045 (2016).
- Spitler, L. G. *et al.* Fast radio burst discovered in the Arecibo pulsar ALFA survey. *Astrophys. J.* **790**, 101 (2014).
- Spitler, L. G. *et al.* A repeating fast radio burst. *Nature* **531**, 202–205 (2016).
- Petroff, E. *et al.* A survey of FRB fields: limits on repeatability. *Mon. Not. R. Astron. Soc.* **454**, 457–462 (2015).
- Chatterjee, S. *et al.* A direct localization of a fast radio burst and its host. *Nature* **541**, 58–61 (2017).
- Tendulkar, S. P. *et al.* The host galaxy and redshift of the repeating fast radio burst FRB 121102. *Astrophys. J.* **834**, L7 (2017).
- Bassa, C. G. *et al.* FRB 121102 is coincident with a star-forming region in its host galaxy. *Astrophys. J.* **843**, L8 (2017).
- Marcote, B. *et al.* The repeating fast radio burst FRB 121102 as seen on millisecond angular scales. *Astrophys. J.* **834**, L8 (2017).
- Bower, G. C., Wright, M. C. H., Falcke, H. & Backer, D. C. Interferometric detection of linear polarization from Sagittarius A* at 230 GHz. *Astrophys. J.* **588**, 331–337 (2003).
- Marrone, D. P., Moran, J. M., Zhao, J.-H. & Rao, R. An unambiguous detection of Faraday rotation in Sagittarius A*. *Astrophys. J.* **654**, L57–L60 (2007).
- Metzger, B. D., Berger, E. & Margalit, B. Millisecond magnetar birth connects frb 121102 to superluminous supernovae and long-duration gamma-ray bursts. *Astrophys. J.* **841**, 14 (2017).
- Piro, A. L. The impact of a supernova remnant on fast radio bursts. *Astrophys. J.* **824**, L32 (2016).
- Scholz, P. *et al.* The repeating fast radio burst FRB 121102: multi-wavelength observations and additional bursts. *Astrophys. J.* **833**, 177 (2016).
- Scholz, P. *et al.* Simultaneous X-ray, gamma-ray, and radio observations of the repeating fast radio burst FRB 121102. *Astrophys. J.* **846**, 80 (2017).
- Oppermann, N. *et al.* Estimating extragalactic Faraday rotation. *Astron. Astrophys.* **575**, A118 (2015).
- Akshori, T., Ryu, D. & Gaensler, B. M. Fast radio bursts as probes of magnetic fields in the intergalactic medium. *Astrophys. J.* **824**, 105 (2016).
- O'Sullivan, S. P. *et al.* Complex Faraday depth structure of active galactic nuclei as revealed by broad-band radio polarimetry. *Mon. Not. R. Astron. Soc.* **421**, 3300–3315 (2012).
- Haverkorn, M. in *Magnetic Fields in Diffuse Media* (eds Lazarian, A. *et al.*) 483–506 (Springer, 2015).
- Hunt, L. K. & Hirashita, H. The size-density relation of extragalactic H II regions. *Astron. Astrophys.* **507**, 1327–1343 (2009).
- Quataert, E., Narayan, R. & Reid, M. J. What is the accretion rate in Sagittarius A*? *Astrophys. J.* **517**, L101–L104 (1999).
- Eatough, R. P. *et al.* A strong magnetic field around the supermassive black hole at the centre of the Galaxy. *Nature* **501**, 391–394 (2013).
- Davidson, K. & Fesen, R. A. Recent developments concerning the Crab Nebula. *Annu. Rev. Astron. Astrophys.* **23**, 119–146 (1985).
- Harvey-Smith, L. *et al.* Faraday rotation of the supernova remnant G296.5+10.0: evidence for a magnetized progenitor wind. *Astrophys. J.* **712**, 1157–1165 (2010).
- Zhang, B. A 'Cosmic comb' model of fast radio bursts. *Astrophys. J.* **836**, L32 (2017).
- Masui, K. *et al.* Dense magnetized plasma associated with a fast radio burst. *Nature* **528**, 523–525 (2015).
- Petroff, E. *et al.* A real-time fast radio burst: polarization detection and multiwavelength follow-up. *Mon. Not. R. Astron. Soc.* **447**, 246–255 (2015).
- Cordes, J. M. *et al.* Lensing of fast radio bursts by plasma structures in host galaxies. *Astrophys. J.* **842**, 35 (2017).
- Manchester, R. N., Hobbs, G. B., Teoh, A. & Hobbs, M. The Australia telescope national facility pulsar catalogue. *Astron. J.* **129**, 1993–2006 (2005).

Acknowledgements We thank the staff of the Arecibo Observatory and the Green Bank Observatory for their help with our observations. We also thank B. Adebahr, L. Connor, G. Desvignes, R. Eatough, R. Fender, M. Haverkorn, A. Karastergiou, R. Morganti, E. Petroff, F. Vieyro and J. Weisberg for suggestions and comments on the manuscript. The Arecibo Observatory is operated by SRI International under a cooperative agreement with the National Science Foundation (AST-1100968), and in alliance with Ana G. Méndez-Universidad Metropolitana and the Universities Space Research Association. The Green Bank Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. Breakthrough Listen (BL) is managed by the Breakthrough Initiatives, sponsored by the Breakthrough Prize Foundation (<http://www.breakthroughinitiatives.org>). The research leading to these results received funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007–2013). J.W.T.H. is a Netherlands Organisation for Scientific Research (NWO) Vidi Fellow and, together with D.M., K.G. and C.G.B., acknowledges funding for this work from ERC Starting Grant DRAGNET under contract number 337062. L.G.S. acknowledges financial support from the ERC Starting Grant BEACON under contract number 279702, as well as the Max Planck Society. A.M.A. is an NWO Veni Fellow. S.C., J.M.C., P.D., T.J.W.L., M.A.M. and S.M.R. are partially supported by the NANOGrav Physics Frontiers Center (NSF award 1430284). V.M.K. holds the Lorne Trotter Chair in Astrophysics & Cosmology and a Canada Research Chair and receives support from an NSERC Discovery Grant and Herzberg Prize, from an R. Howard Webster Foundation Fellowship from the Canadian Institute for Advanced Research (CIFAR), and from the FRQNT Centre de Recherche en Astrophysique du Québec. C.J.L. acknowledges support from NSF award 1611606. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. B.M. acknowledges support from the Spanish Ministerio de Economía y Competitividad (MINECO) under grants AYA2016-76012-C3-1-P and MDM-2014-0369 of ICCUB (Unidad de Excelencia 'María de Maeztu'). S.M.R. is a CIFAR Senior Fellow. P.S. holds a Covington Fellowship at DRAC.

Author Contributions A.S. led the development of the Arecibo observing functionality used here and discovered the first bursts near 4.5 GHz. L.G.S. is Principal Investigator of the Arecibo monitoring campaign. D.M. discovered the rotation measure and analysed the burst properties in detail. K.G. searched all Arecibo datasets near 4.5 GHz for bursts. J.W.T.H. led the discussion on the interpretation of the results and writing of the manuscript. A.M.A. guided the development of the rotation measure fitting code. G.H.H. and C.S. performed the rotation measure synthesis and deconvolution analysis. G.C.B., S.C., J.M.C., V.G., V.M.K., C.J.L., M.A.M. and D.M. also contributed to the writing of the manuscript and analysis. V.G. observed, searched and detected bursts from the GBT at 6.5 GHz as a part of the BL monitoring campaign of known fast radio bursts. A.P.V.S. is the Principal Investigator of the BL project. C.B. helped with the polarization calibration of the test pulsar. G.H. wrote a code to splice raw voltages across computer nodes. All other co-authors contributed to the interpretation of the analysis results and to the final version of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to J.W.T.H. (J.W.T.Hessels@uva.nl).

Reviewer Information Nature thanks H. Falcke and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

The analyses described here were based on the PRESTO³¹, PSRCHIVE³² and DSPSR³³ pulsar software suites, as well as custom-written Python scripts for linking utilities into reduction pipelines, fitting the data and plotting.

Observations and burst search. *Arecibo.* We made the observations using the Arecibo ‘C-band’ receiver (dual linear receptors) in the frequency range 4.1–4.9 GHz and the Puerto-Rican Ultimate Pulsar Processing Instrument (PUPPI) backend recorder. The full list of observations is reported in Extended Data Table 1. We operated PUPPI in its ‘coherent search’ mode, which produced 10.24 μ s samples and 512×1.56 MHz frequency channels, each coherently dedispersed to dispersion measure 557.0 pc cm⁻³. Coherent dedispersion within each 1.56-MHz channel means that the intra-channel dispersive smearing is smaller than 2 μ s even if the dispersion measure of the burst is 10 pc cm⁻³ higher or lower than the fiducial value of 557.0 pc cm⁻³ used in the PUPPI recording. The raw PUPPI data also provide auto- and cross-correlations of the two linear polarizations, which can be converted to the Stokes parameters I , Q , U and V during post-processing. Before each observation, we performed a test scan on a known pulsar (PSR B0525+21) and a noise diode calibration scan (for polarimetric calibration).

Dedispersed time series with dispersion measure from 461 pc cm⁻³ to 661 pc cm⁻³ were searched using trial steps of 1 pc cm⁻³ and the PRESTO routine *single_pulse_search.py*, which applies a matched-filter technique to look for bursts with durations between 81.92 μ s and 24,576 μ s (for any putative burst that only has a single peak with width below 81.92 μ s, the sensitivity will be degraded by a factor of a few at most). The resulting data points (dispersion measure, time, signal-to-noise ratio) were grouped into plausible astrophysical burst candidates using a custom sifting algorithm and then a dynamic spectrum of each candidate was plotted for manual inspection and grading. We found 16 bursts of astrophysical origin and used the DSPSR package to form full-resolution, full-polarization PSRCHIVE ‘archive’ format files for each burst.

Green Bank Telescope. On 26 August 2017, we observed FRB 121102 using the GBT ‘C-band’ receiver (4–8 GHz, with dual linear receptors) as part of a programme of monitoring known FRB positions. Observations were conducted with the Breakthrough Listen Digital Backend³⁴, which allowed recording of baseband voltage data across the entire nominal 4-GHz bandwidth of the selected receiver. Scans of a noise diode calibration, of the flux calibrator 3C161 and of the bright pulsar PSR B0329+54 supplemented the observations.

In post-processing, a total-intensity, low-resolution filterbank data product was searched for bursts with dispersion measure in the range 500–600 pc cm⁻³ using trial dispersion measure values in steps of 0.1 pc cm⁻³ and a search package implemented on an accelerated graphics processing unit to perform incoherent dedispersion³⁵. We detected³⁶ 15 bursts with signal-to-noise ratio higher than 10. Here we present the properties of just the two brightest GBT bursts in order to confirm the large rotation measure observed by Arecibo and to quantify its variation in time. A detailed analysis of all GBT detections is in progress (V.G. *et al.*, manuscript in preparation). A section of raw voltage data (1.5 s) around each detected burst was extracted and coherently dedispersed to a nominal dispersion measure of 557.91 pc cm⁻³ using the DSPSR package. The final PSRFITS format data products have time and spectral resolutions of 10.24 μ s and 183 kHz, respectively.

Data analysis. *Calculation of burst rotation measures.* We calibrated the burst ‘archives’ using the PSRCHIVE utility *pac* in ‘SingleAxis’ mode. This calibration strategy uses observations of a locally generated calibration signal (pulsed noise diode) to correct the relative gain and phase difference between the two polarization channels, under the assumption that the noise source emits equal power and has zero intrinsic phase difference in the two polarization channels. This calibration scheme does not correct for cross-coupling or leakage between the polarizations. While leakage must be present at some level, the high polarization fraction, complete lack of circular polarization, and consistency of the test pulsar observations with previous work give us confidence that calibration issues are not a substantial source of error for the rotation measure determination. In addition, the flux density of GBT observations was calibrated using the flux calibrator.

We initially performed a brute-force search for peaks in the linear polarization fraction (Extended Data Fig. 3) and discovered that $RM_{\text{obs}} \approx +10^5$ rad m⁻² in the Arecibo data. Each burst was corrected for Faraday rotation using the best-fit rotation measure value for that burst. Residual variations in the resulting polarization angle $PA(\lambda)$ were used to refine the initial values by fitting

$$PA(\lambda) = RM\lambda^2 + PA_{\infty} \quad (1)$$

and then

$$\hat{L} = \exp[i2(RM\lambda^2 + PA_{\infty})] \quad (2)$$

where \hat{L} is the unit vector of the linear polarization. We used equation (2) to fit the whole sample of bursts, imposing a different rotation measure per day and a different PA_{∞} per telescope. The results of these fits are reported in Table 1 and an example is shown in Fig. 2.

By applying the optimal rotation measure value to each burst, we produced polarimetric profiles showing that each burst is almost 100% linearly polarized, after accounting for the finite widths of the PUPPI frequency channels (Fig. 1; Extended Data Fig. 2). In fact, the measured Arecibo bursts are depolarized to about 98%, in agreement with an uncorrected intra-channel Faraday rotation of

$$\Delta\theta = \frac{RM_{\text{obs}}c^2\Delta\nu}{\nu_c^3} \quad (3)$$

where c is the speed of light, $\Delta\nu$ is the channel width and ν_c is the central channel observing frequency. At 4.5 GHz, this corresponds to about 9°, and the depolarization fraction is

$$f_{\text{depol}} = 1 - \frac{\sin(2\Delta\theta)}{2\Delta\theta} = 1.6\% \quad (4)$$

We supplemented the above analysis with a combination of the RM Synthesis method and the deconvolution procedure RMCLEAN (for example, Extended Data Fig. 4). Ensuring the presence of minimal Faraday complexity is possible by integrating across the full bandwidth and taking advantage of a Fourier transform relation between the observed $L(\lambda^2)$ values and the Faraday spectrum (the polarized brightness as a function of rotation measure). This approach is known as RM Synthesis³⁷ and can be coupled with RMCLEAN to estimate the intrinsic Faraday spectrum³⁸. Although RM Synthesis and RMCLEAN can have difficulty in properly reconstructing the intrinsic Faraday spectrum under certain circumstances, the spread of clean components is a reliable indicator of spectra that contain more than a single Faraday-unresolved source³⁹.

At each observed frequency, we integrated the Stokes parameters Q and U across the pulse width and normalized using the Stokes I profile. Owing to the normalization, we used only frequency bins that had a Stokes I signal-to-noise ratio of at least 5. We computed a deconvolved Faraday spectrum for each burst separately on a highly oversampled rotation measure axis (RM sampling $\delta RM \approx 10^{-4}$ of the nominal full-width at half-maximum of the rotation measure resolution element). We used a relatively small gain parameter (0.02) and terminated the deconvolution when the peak of the residual decreased to 2σ above the mean. The algorithm typically required 50–80 iterations to converge. This combination of settings permits us to carefully consider the cumulative distribution of RMCLEAN components along the rotation measure axis and thus constrain the intrinsic width of the polarized emission to below about 0.1% of the typical rotation measure uncertainty. We found that this value scales with δRM because the peak of the Faraday spectrum rarely lands precisely on an individual pixel. To a high degree of confidence, there is evidence neither of emission at more than one rotation measure value, nor of a broadened (‘Faraday-thick’) emitting region; we therefore forgo more complicated fitting of the Q and U parameters¹⁹. The results of this analysis, shown in Extended Data Table 2, are consistent with the simplified fitting results described above.

Calculation of burst properties. As in previous studies^{5,15}, a search for periodicity in the burst arrival times remains inconclusive. Determining the exact dispersion measures of the bursts is complicated by their changing morphology with radio frequency^{5,15}. Measuring the dispersion measure based on maximizing the peak signal-to-noise ratio of the burst often leads to blurring of the burst structure and, in the case of FRB 121102, an overestimation of the dispersion measure. We have thus chosen to display all bursts dedispersed to the same nominal dispersion measure of burst 6 (Fig. 1 and Extended Data Fig. 1). Taking advantage of the narrowness of burst 6, we estimated its optimal dispersion measure by minimizing its width at different dispersion measure trials. We measured burst widths at half the maximum of the peak value by fitting them with von Mises functions using the PSRCHIVE routine *paas* (Table 1). These widths correspond to the burst envelope in the case of multi-component bursts.

The flux densities of the Arecibo bursts were estimated using the radiometer equation to calculate the equivalent root-mean-square flux density of the noise

$$\sigma_{\text{noise}} = \frac{T_{\text{sys}}}{G\sqrt{2Bt_{\text{int}}}} \quad (5)$$

where $T_{\text{sys}} \approx 30$ K and $G \approx 7$ K Jy⁻¹ are the system temperature and gain of the receiver, respectively, $B = 800$ MHz is the observing bandwidth and $t_{\text{int}} = 10.24 \mu$ s is the sampling time. The GBT observations were calibrated using a flux calibrator, as discussed earlier. Because of the complicated spectra of the bursts, we quote average values across the frequency band (Table 1).

The dynamic spectra of the bursts in Extended Data Fig. 1 show narrow-band striations that are consistent with diffractive interstellar scintillations caused by turbulent plasma in the Milky Way. The autocorrelation functions of the burst spectra show three features: a very narrow feature from radiometer noise, a narrow but resolved feature corresponding to the striations, and a broad feature related to the extent of the burst across the frequency band. The striation feature has a half-width that varies between about 2 MHz and 5 MHz from burst to burst and is comparable to the scintillation bandwidth expected from the Milky Way in the direction of FRB 121102. The NE2001 electron density model⁴⁰ provides an estimate $\tau \approx 16 \mu\text{s}$ for the pulse broadening time at 1 GHz. This predicts a scintillation bandwidth of about $\nu^{4.4}/(2\pi\tau)$ that ranges from 5 MHz to 11 MHz across the 4.1–4.9 GHz band. We conclude that the measured autocorrelation functions and the NE2001 model prediction are consistent to within their uncertainties and that the narrow striations are due to Galactic scintillations.

A model for the rotation measure and scattering measure of FRB 121102.

Rotation measure constraints. The measured $\text{RM}_{\text{obs}} \approx +1 \times 10^5 \text{ rad m}^{-2}$ implies a source frame value

$$\text{RM}_{\text{src}} = (1+z)^2 \text{RM}_{\text{obs}} \approx +1.4 \times 10^5 \text{ rad m}^{-2} \quad (6)$$

We can use the previously estimated⁸ $\text{DM}_{\text{host}} \approx 70\text{--}270 \text{ pc cm}^{-3}$ (in the source frame) and RM_{src} to constrain the properties of the region in which the Faraday rotation occurs. In the absence of other information, we can set a constraint on the average magnetic field along the line of sight in the Faraday region using the ratio

$$B_{\parallel} = \frac{\text{RM}_{\text{src}}}{0.81 \text{DM}_{\text{host}}} = [0.6 \text{ mG}, 2.4 \text{ mG}] \quad (7)$$

If only a small portion of the total dispersion measure of FRB 121102 is from the highly magnetized region, the field could be much higher.

Scattering measure constraints. The best constraint on pulse broadening comes from the measurement of the scintillation (diffraction) bandwidth of $\Delta\nu_d \approx 5 \text{ MHz}$ at 4.5 GHz (see above). This implies a pulse broadening time at 1 GHz of

$$\tau(1 \text{ GHz}) \approx (2\pi\Delta\nu_d)^{-1} \times (4.5 \text{ GHz}/1 \text{ GHz})^{4.4} = 24 \mu\text{s} \quad (8)$$

This scattering time is consistent with that expected from the Milky Way using the NE2001 model⁴⁰ and therefore is an upper bound on any contribution from the host galaxy. Compared to scattering in the Milky Way, this upper bound is below the mean trend for any of the plausible values of DM_{host} , especially when the correction from spherical to plane waves is taken into account⁴¹.

Compared with the observer frame, the ratio $\tau/\text{DM}_{\text{host}}$ is a factor of $(1+z)^2 = 1.42$ larger in the source frame but that is still far from sufficient to account for the apparent scattering deficit with respect to the Galactic τ/DM ratio. Given the apparent extreme conditions of the plasma in the host galaxy, it would not be surprising if its turbulence properties caused a scattering deficit. For example, scattering is reduced if the inner scale is comparable to or larger than the Fresnel scale owing to either a large magnetic field or a high temperature.

Constraints on the properties of the Faraday region. Comparison of the magnetic field and thermal energy densities enables us to constrain the electron density and temperature and the length scale of the region responsible for the observed Faraday rotation. We parametrize this relation with

$$\beta \frac{B^2}{8\pi} = 2n_e k_B T_e \quad (9)$$

where β is a scaling factor, B is the magnetic field strength and k_B is the Boltzmann constant. This assumes a 100% ionized gas of pure hydrogen with temperature equilibration between protons and electrons. Under equipartition, $\beta = 1$. In more densely magnetized regions, $\beta \ll 1$. Field reversals will reduce the total rotation measure, requiring a lower value of β in order to match constraints. The absence of free-free absorption at a frequency of about 1 GHz sets an additional constraint on the permitted parameter space.

In Extended Data Fig. 6, we explore a range of physical environments. We consider a lower limit, $\text{DM} = 1 \text{ pc cm}^{-3}$, on the dispersion measure that is smaller than the previously estimated⁸ $\text{DM}_{\text{host}} \approx 70\text{--}270 \text{ pc cm}^{-3}$ because it is possible that not all of the dispersion measure originates from the Faraday region. Galactic H II regions typically have $|\text{RM}|$ smaller than about $3 \times 10^2 \text{ rad m}^{-2}$ and weak magnetic fields⁴² with β greater than about 1, although calculations suggest that it is possible for H II regions to achieve high rotation measures under some circumstances⁴³. The parameter space for a typical H II-region plasma at $T_e = 10^4 \text{ K}$ is almost entirely excluded, and many possible H II region sizes and densities²¹ are incompatible with the DM_{host} constraints. At higher T_e , wide ranges of the parameter space are permitted. In the case of equipartition, we have explicit unique solutions. For $T_e = 10^6 \text{ K}$, we find a density of $n_e \approx 10^2 \text{ cm}^{-3}$ on a length scale $L_{\text{RM}} \approx 1 \text{ pc}$,

comparable to the upper limit of the size of the persistent source. Higher-temperature gas ($T_e = 10^8 \text{ K}$) can be extended to $L_{\text{RM}} \approx 100 \text{ pc}$. For both of these solutions, the characteristic magnetic field strength is about 1 mG.

The large rotation measure of FRB 121102 is similar to those seen towards massive black holes; notably, $\text{RM} \approx -5 \times 10^5 \text{ rad m}^{-2}$ is measured near Sagittarius A*, the Milky Way's central black hole, and probes scales below 10^4 Schwarzschild radii (about 0.001 pc)^{11,12}. The constraints on n_e , T_e and L_{RM} are also consistent with the environment around Sagittarius A* (Extended Data Fig. 6). The high rotation measure towards the Galactic Centre magnetar PSR J1745–2900 (Fig. 3), $\text{RM} = -7 \times 10^4 \text{ rad m}^{-2}$, at a projected distance^{23,44} of about 0.1 pc from Sagittarius A*, is evidence of a dynamically organized magnetic field around Sagittarius A* that extends to the distance of the magnetar²³. Notably, radio monitoring of PSR J1745–2900 for about 4.5 years has shown a decrease of around 5% in the magnitude of the observed rotation measure, while the dispersion measure remained constant at a level of about 1% (Desvignes, G. *et al.*, manuscript in preparation). This suggests large fluctuations in magnetic field strength in the Galactic Centre on scales of roughly 10^{-5} pc .

The high rotation measure and the rich variety of other phenomena^{4,5,7–10,15,16} displayed by the FRB 121102 system suggest that the persistent radio counterpart to FRB 121102 could represent emission from an accreting massive black hole, with the surrounding star formation representing a circum-black-hole starburst. Given the mass of the host galaxy and typical scaling relationships⁴⁵, the mass of the black hole would be about $10^4\text{--}10^6$ solar masses (M_{\odot}). The observed radio brightness and compactness of the source, as well as the optical and X-ray non-detections^{8,10,16}, are compatible with such a black hole and an inefficient accretion state (about $10^{-6} L_{\text{Edd}}\text{--}10^{-4} L_{\text{Edd}}$, where L_{Edd} is the Eddington luminosity).

While models considering the presence of only a massive black hole have been proposed⁴⁶, there is no observational precedent for microsecond bursts created in such environments. Rather, the FRB 121102 bursts themselves could arise from a neutron star, perhaps highly magnetized and rapidly spinning, near an accreting massive black hole. The proximity of PSR J1745–2900 to Sagittarius A* demonstrates that such a combination is possible. In this model, the black hole is responsible for the observed persistent source, whereas the bursts are created in the magnetosphere of the nearby neutron star⁴⁷.

Alternatively, the association of FRB 121102 with a persistent radio source has been used to argue that the radio bursts are produced by a young magnetar powering a luminous wind nebula^{13,48}. This model is not well motivated by Galactic examples, since the most luminous (non-magnetar-powered) Galactic pulsar wind nebula is 500,000 times less luminous than the persistent source that is coincident with FRB 121102, and Galactic magnetars have no detectable persistent radio wind nebulae^{49,50}. Also, although giant flares from magnetars can produce relativistic outflows⁵¹, an upper limit of the rotation measure from one such outburst⁵² is four orders of magnitude below that observed for FRB 121102.

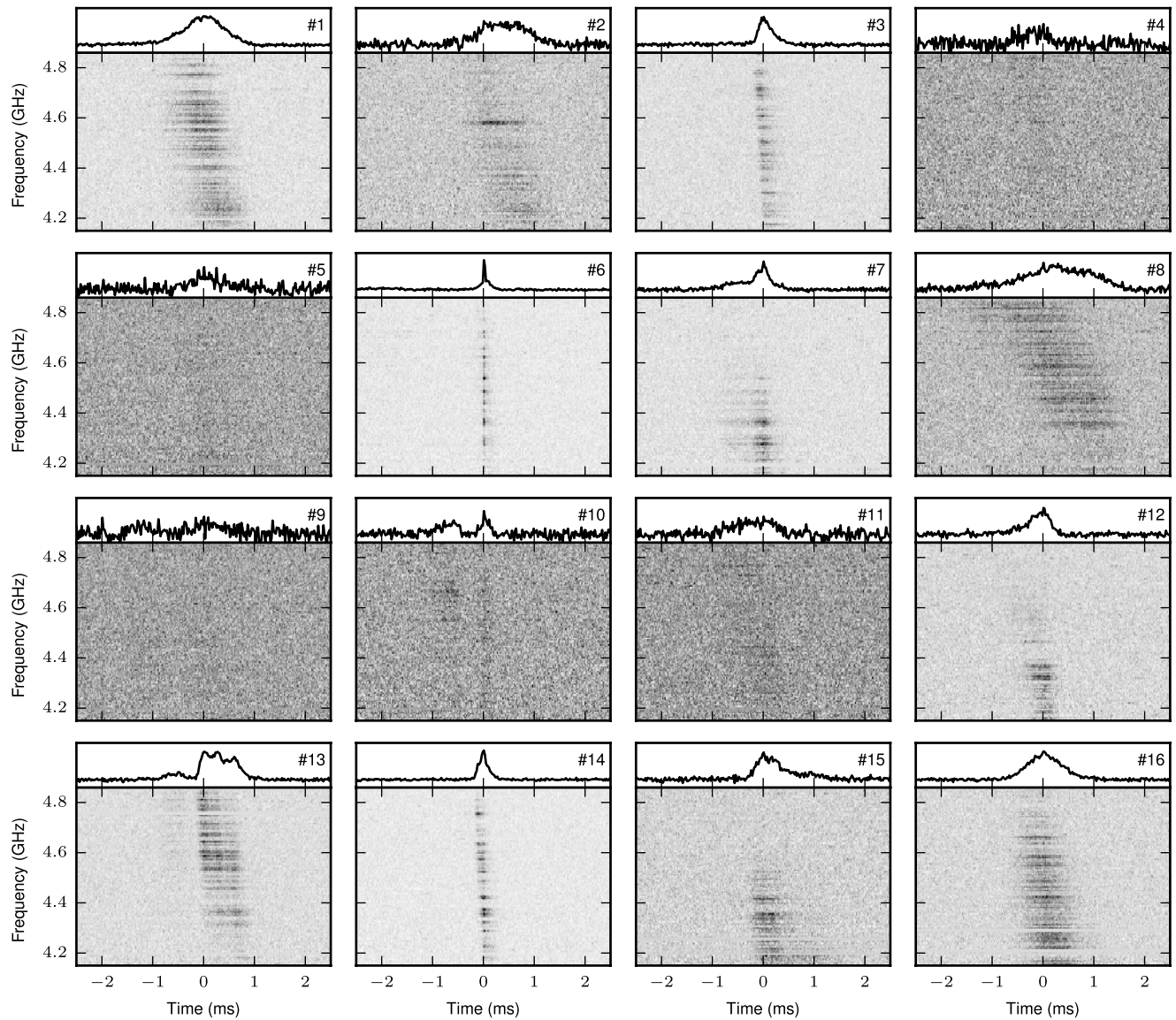
Nonetheless, under the millisecond magnetar model, the properties of the persistent source constrain the age of the putative magnetar to between several years and several decades, with a spin-down luminosity of 10^8 to 10^{12} times higher than any local analogue¹³. Furthermore, the millisecond magnetar model predicts that the nebula magnetic field strength scales with the integrated spin-down luminosity of the magnetar^{13,48}. Extended Data Fig. 6 shows a range of sizes, densities and temperatures for the Faraday-rotating medium that are consistent with Crab-like pulsar wind nebulae, known supernova remnants and a simple model for swept-up supernova ejecta.

Data availability. The calibrated burst data are available upon request from the corresponding author.

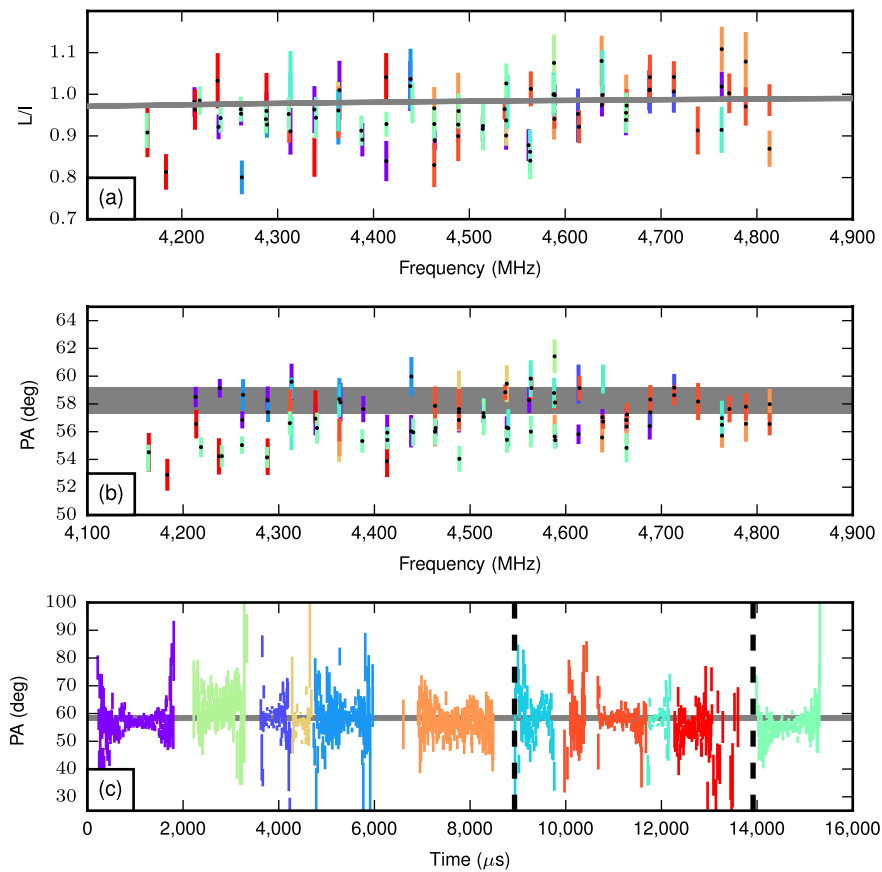
Code availability. The codes used to analyse the data are available at the following sites: PRESTO (<https://github.com/scotttransom/presto>), PSRCHIVE (<http://psrchive.sourceforge.net>) and DSPSR (<http://dspsr.sourceforge.net>).

1. Ransom, S. M. *New Search Techniques for Binary Pulsars*. PhD thesis, Harvard Univ. (2001); <http://adsabs.harvard.edu/abs/2001PhDT.....123R>.
2. van Straten, W., Demorest, P. & Osłowski, S. Pulsar data analysis with PSRCHIVE. *Astron. Res. Technol.* **9**, 237–256 (2012).
3. van Straten, W. & Bailes, M. DSPSR: digital signal processing software for pulsar astronomy. *Publ. Astron. Soc. Aust.* **28**, 1–14 (2011).
4. MacMahon, D. H. E. *et al.* The breakthrough listen search for intelligent life: a wideband data recorder system for the Robert C. Byrd green bank telescope. *Publ. Astron. Soc. Pac.* (in the press); preprint at <https://arxiv.org/abs/1707.06024> (2017).
5. Barsdell, B. R., Bailes, M., Barnes, D. G. & Fluke, C. J. Accelerating incoherent dedispersion. *Mon. Not. R. Astron. Soc.* **422**, 379–392 (2012).
6. Gajjar, V. *et al.* FRB 121102: detection at 4–8 GHz band with Breakthrough Listen backend at Green Bank. *Astron. Telegr.* **10675** (2017).
7. Brentjens, M. A. & de Bruyn, A. G. Faraday rotation measure synthesis. *Astron. Astrophys.* **441**, 1217–1228 (2005).

38. Heald, G., Braun, R. & Edmonds, R. The Westerbork SINGS survey. II. Polarization, Faraday rotation, and magnetic fields. *Astron. Astrophys.* **503**, 409–435 (2009).
39. Anderson, C. S., Gaensler, B. M., Feain, I. J. & Franzen, T. M. O. Broadband radio polarimetry and Faraday rotation of 563 extragalactic radio sources. *Astrophys. J.* **815**, 49 (2015); erratum 820, 144 (2016).
40. Cordes, J. M. & Lazio, T. J. W. NE2001.I. A new model for the galactic distribution of free electrons and its fluctuations. Preprint at <http://arxiv.org/abs/astro-ph/0207156> (2002).
41. Cordes, J. M., Wharton, R. S., Spitler, L. G., Chatterjee, S. & Wasserman, I. Radio wave propagation and the provenance of fast radio bursts. Preprint at <https://arxiv.org/abs/1605.05890> (2016).
42. Harvey-Smith, L., Madsen, G. J. & Gaensler, B. M. Magnetic fields in large-diameter H II regions revealed by the Faraday rotation of compact extragalactic radio sources. *Astrophys. J.* **736**, 83 (2011).
43. Sicheneder, E. & Dexter, J. A single H II region model of the strong interstellar scattering towards Sgr A*. *Mon. Not. R. Astron. Soc.* **467**, 3642–3647 (2017).
44. Shannon, R. M. & Johnston, S. Radio properties of the magnetar near Sagittarius A* from observations with the Australia Telescope Compact Array. *Mon. Not. R. Astron. Soc.* **435**, L29–L32 (2013).
45. Reines, A. E. & Volonteri, M. Relations between central black hole mass and total galaxy stellar mass in the local universe. *Astrophys. J.* **813**, 82 (2015).
46. Vieyro, F. L., Romero, G. E., Bosch-Ramon, V., Marcote, B. & del Valle, M. V. A model for the repeating FRB 121102 in the AGN scenario. *Astron. Astrophys.* **602**, A64 (2017).
47. Pen, U.-L. & Connor, L. Local circumnuclear magnetar solution to extragalactic fast radio bursts. *Astrophys. J.* **807**, 179 (2015).
48. Kashiwama, K. & Murase, K. Testing the young neutron star scenario with persistent radio emission associated with FRB 121102. *Astrophys. J.* **839**, L3 (2017).
49. Hester, J. J. The Crab Nebula: an astrophysical chimera. *Annu. Rev. Astron. Astrophys.* **46**, 127–155 (2008).
50. Reynolds, S. P. *et al.* Pulsar-wind nebulae and magnetar outflows: observations at radio, x-ray, and gamma-ray wavelengths. *Space Sci. Rev.* **207**, 175–234 (2017).
51. Frail, D. A., Kulkarni, S. R. & Bloom, J. S. An outburst of relativistic particles from the soft γ -ray repeater SGR1900+14. *Nature* **398**, 127–129 (1999).
52. Gaensler, B. M. *et al.* An expanding radio nebula produced by a giant flare from the magnetar SGR 1806-20. *Nature* **434**, 1104–1106 (2005).
53. McKee, C. F. & Truelove, J. K. Explosions in the interstellar medium. *Phys. Rep.* **256**, 157–172 (1995).
54. Orlando, S., Miceli, M., Pumo, M. L. & Bocchino, F. Modeling SNR Cassiopeia A from the supernova explosion to its current age: the role of post-explosion anisotropies of ejecta. *Astrophys. J.* **822**, 22 (2016).
55. McCray, R. & Fransson, C. The remnant of supernova 1987A. *Annu. Rev. Astron. Astrophys.* **54**, 19–52 (2016).

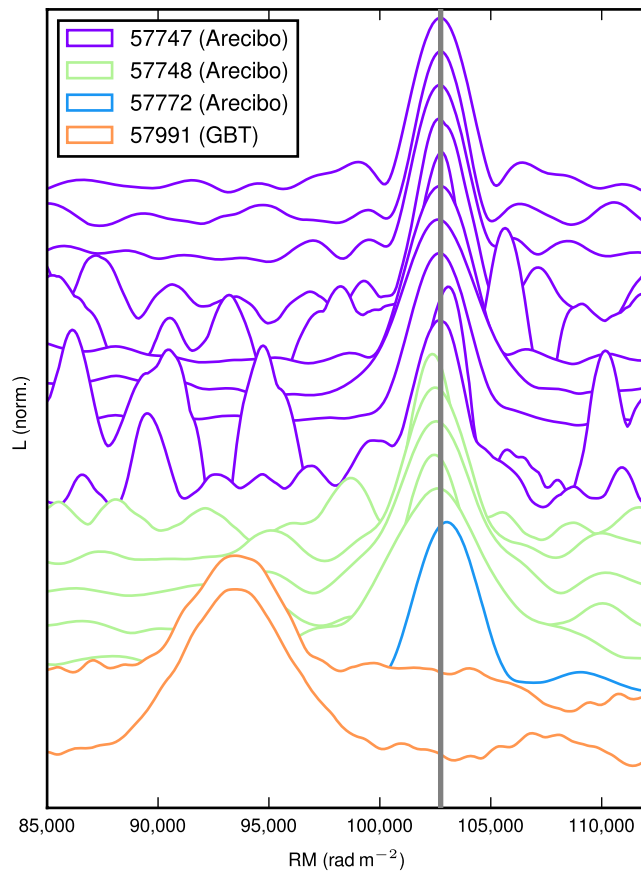


Extended Data Figure 1 | Pulse profiles and spectra of 16 Arecibo bursts. The bursts are dedispersed to $DM = 559.7 \text{ pc cm}^{-3}$ (which minimizes the width of burst 6) and plotted with time and frequency resolutions of $20.48 \mu\text{s}$ and 6.24 MHz , respectively.

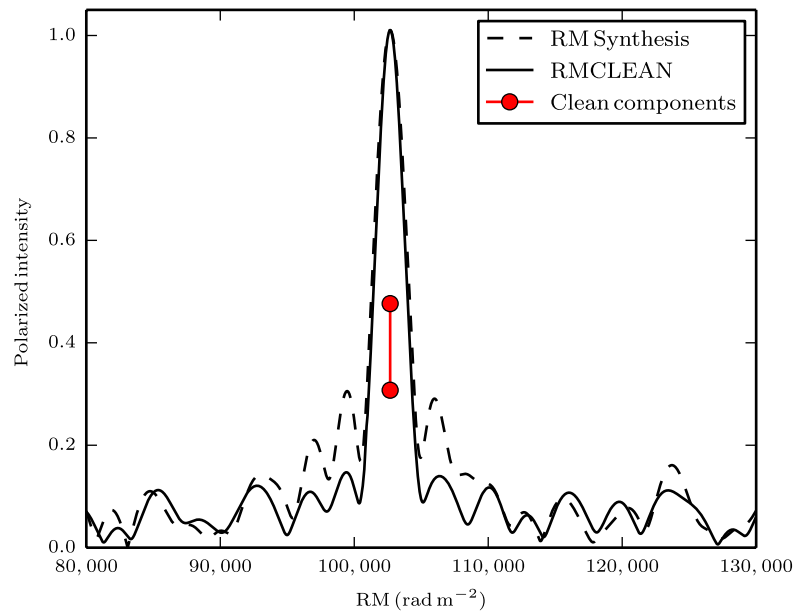


Extended Data Figure 2 | Polarimetric properties of the 11 brightest bursts detected by Arecibo. **a**, Linear polarization fraction of the bursts as a function of frequency. The solid line shows the theoretical depolarization due to intra-channel Faraday rotation, calculated using equations (3) and (4). **b**, PA_{∞} as a function of frequency. Values in **a** and **b** are averaged over 16 consecutive channels. **c**, PA_{∞} as a function of time. A time offset

is applied to each burst in order to show them consecutively. Vertical dashed lines divide different observing sessions. All values in this figure have been corrected for the rotation measure, which was calculated with a global fit. Grey regions in **b** and **c** indicate the 1σ uncertainty around the polarization angle determined from the global fit.

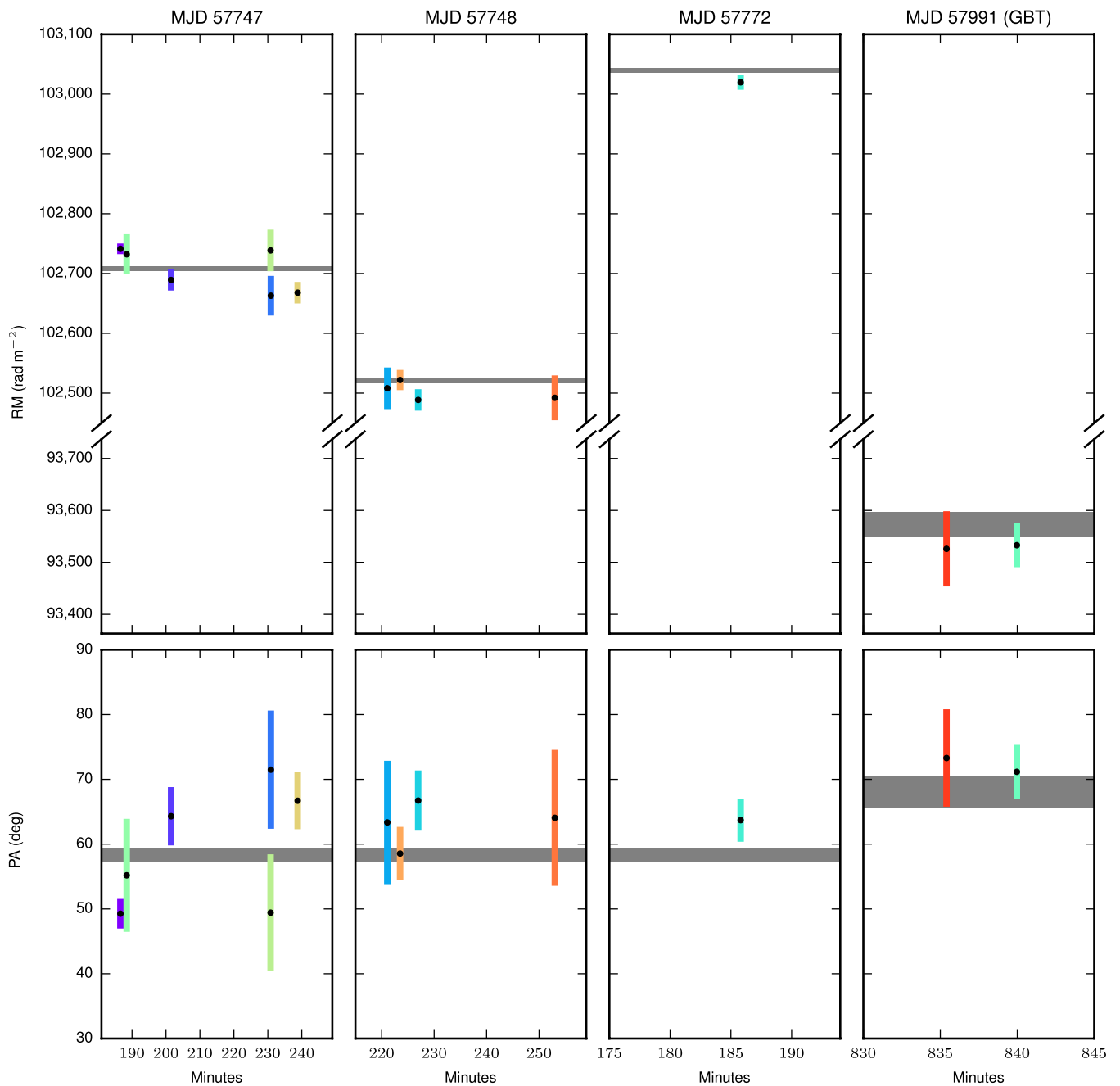


Extended Data Figure 3 | Linear polarization fraction of the bursts as a function of rotation measure. Different colours represent different observing sessions (see key). The grey line indicates the average rotation measure that yields the largest polarization fraction in the first observing session.



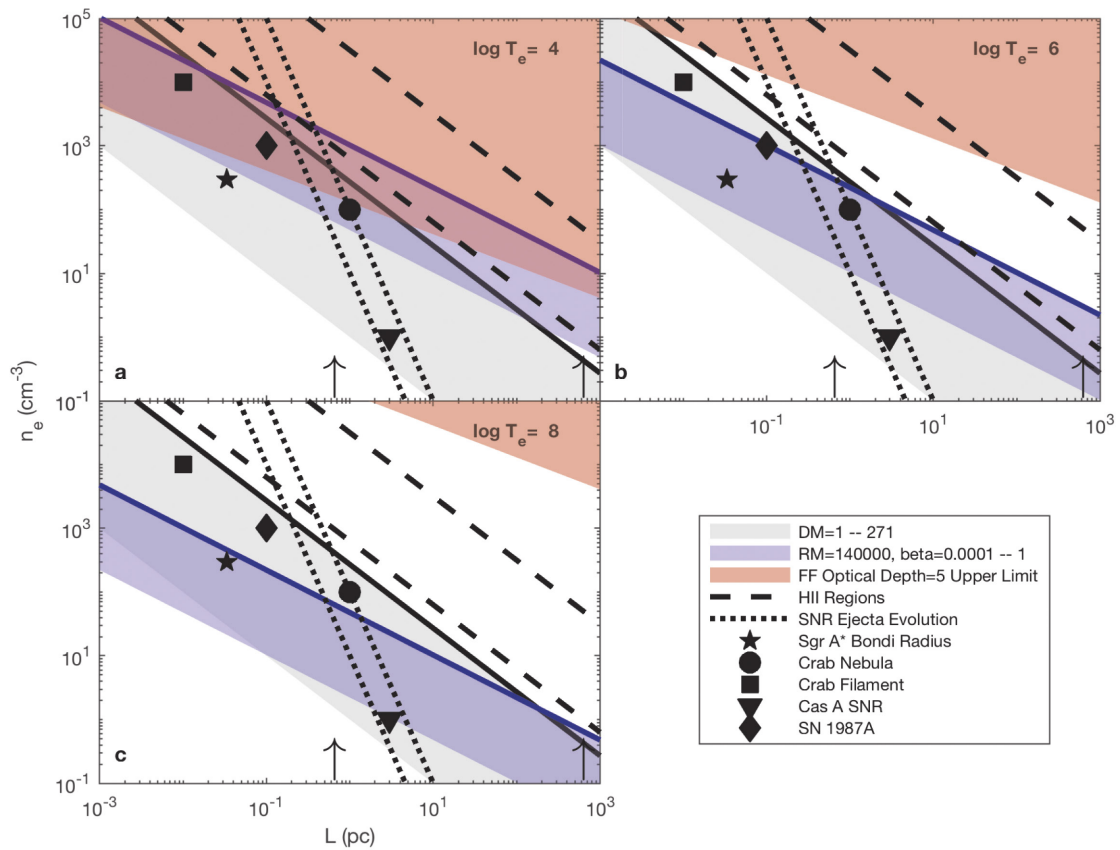
Extended Data Figure 4 | Example of RM Synthesis and RMCLEAN results for burst 8. The relevant rotation measure range is shown for burst 8, after analysis with RM Synthesis (dashed line) and RMCLEAN (solid line), as described in the main text. Only two clean components (red circles) were required to reach convergence in the deconvolution

algorithm (at 102,679.5 rad m^{-2} and 102,679.75 rad m^{-2} ; compare with the peak of the final deconvolved Faraday spectrum at 102,679.65 rad m^{-2}). For all bursts, the RM Synthesis and RMCLEAN steps demonstrate an extremely thin and single-peaked Faraday spectrum.



Extended Data Figure 5 | Rotation measure and PA_{∞} values of different bursts. Coloured, 1σ error bars represent individual bursts, with central values highlighted by black dots. Horizontal grey regions are values

obtained from a global fit. MJD, modified Julian date. Values used in the figure are reported in Table 1.



Extended Data Figure 6 | Physical constraints from source parameters. **a–c**, Parameter space for the electron density (n_e) and length scale (L_{RM}) of the Faraday region for three different temperature regimes, $T_e = 10^4$ K (**a**), 10^6 K (**b**) and 10^8 K (**c**). The shaded red region indicates the parameter space excluded because of optical depth considerations (optical depth from free–free absorption $\tau_{\text{ff}} > 5$). The solid black line indicates the maximum DM_{host} permitted, while the shaded grey region shows the dispersion measure down to 1 pc cm^{-3} . The solid blue line denotes RM_{src} . The shaded blue region shows the range $10^{-4} \leq \beta \leq 1$. The intersection of grey and blue regions outside of the red region is physically permitted. The arrows indicate the upper limits on the sizes of the persistent source

(left) and the star-forming region (right), respectively^{8,10}. The parallel dashed lines represent fits to a range of Galactic and extragalactic H II regions²¹. The parallel dotted lines represent the evolution of $1M_{\odot}$ and $10M_{\odot}$ of ejecta in up to 1,000 years at a velocity of 10^4 km s^{-1} in the blast-wave phase following a supernova⁵³. The filled downward-pointing triangles and diamonds correspond to the supernova remnants Cas A and SN 1987A, respectively^{54,55}. The filled circles represent the mean density and diameter of the Crab Nebula, whereas the filled squares represent the characteristic density and length scale of a dense filament in the Crab Nebula²⁴. The stars indicate the density of Sagittarius A* at the Bondi radius²².

Extended Data Table 1 | List of 4.5-GHz Arecibo observations used in this study

Start (MJD)	Duration (s)	# bursts
57717.2018171	4146	0
57717.2500000	2378	0
57747.1172685	6703	10
57748.1141435	6806	5
57772.0590625	5901	1
57806.9996759	2700	0
57813.9342940	5588	0
57821.9134144	2893	0
57858.8624769	2350	0
57865.8491782	1904	0
57872.8160417	2919	0
57900.7106597	2779	0

Extended Data Table 2 | Results of analysis with RM Synthesis and RMCLEAN

Burst	RM (rad m ⁻²)	RM disp (rad m ⁻²)
1	+102805 ± 37	< 0.12
2	+102685 ± 70	< 0.05
3	+102667 ± 37	< 0.12
6	+102642 ± 73	< 0.11
7	+102643 ± 105	< 0.04
8	+102680 ± 43	< 0.12
12	+102585 ± 67	< 0.02
13	+102484 ± 53	0
14	+102440 ± 51	0
15	+102701 ± 211	< 0.05
16	+102986 ± 27	< 0.10
GBT -1	+93572 ± 2885	0
GBT -2	+93523 ± 237	0

Rotation measures were determined by fitting a quadratic function to the peak of the deconvolved Faraday spectrum. Rotation measure uncertainties were determined by dividing the nominal full-width at half-maximum of the rotation measure resolution element by twice the signal-to-noise ratio at the peak of the rotation measure spectrum. RMdisp is the second moment (dispersion) of the RMCLEAN clean components discovered during the Faraday spectrum deconvolution. Upper limits indicate that the value scales with the rotation measure pixel size; a zero value means that all clean components fell within the same pixel and indicates a Faraday spectrum that is indistinguishable from being infinitely thin.

A rapid decrease in the rotation rate of comet 41P/Tuttle–Giacobini–Kresák

Dennis Bodewits¹, Tony L. Farnham¹, Michael S. P. Kelley¹ & Matthew M. Knight¹

Cometary outgassing can produce torques that change the spin state of the cometary nucleus, which in turn influences the evolution and lifetime of the comet^{1,2}. If these torques increase the rate of rotation to the extent that centripetal forces exceed the material strength of the nucleus, the comet can fragment³. Torques that slow down the rotation can cause the spin state to become unstable, but if the torques persist the nucleus can eventually reorient itself and the rotation rate can increase again⁴. Simulations predict that most comets go through a short phase of rapid changes in spin state, after which changes occur gradually over longer times⁵. Here we report observations of comet 41P/Tuttle–Giacobini–Kresák during its close approach to Earth (0.142 astronomical units, approximately 21 million kilometres, on 1 April 2017) that reveal a rapid decrease in rotation rate. Between March and May 2017, the apparent rotation period of the nucleus increased from 20 hours to more than 46 hours—a rate of change of more than an order of magnitude larger than has hitherto been measured. This phenomenon must have been caused by the gas emission from the comet aligning in such a way that it produced an anomalously strong torque that slowed the spin rate of the nucleus. The behaviour of comet 41P/Tuttle–Giacobini–Kresák suggests that it is in a distinct evolutionary state and that its rotation may be approaching the point of instability.

The combination of close approach, brightness and large solar elongation of comet 41P/Tuttle–Giacobini–Kresák (hereafter 41P) made it the target of observations worldwide for several months. We report results from our observations of comet 41P obtained in March 2017 using the Large Monolithic Imager on Lowell Observatory's 4.3-m Discovery Channel Telescope (DCT) and in May 2017 using the UltraViolet–Optical Telescope (UVOT) on board the Earth-orbiting Swift Gamma Ray Burst Mission⁶ (Extended Data Table 1).

We used comet-specific narrowband filters⁷ on the DCT to capture the emission of cyanogen gas. Cyanogen coma structures have been used to infer rotational properties of otherwise unobservable comet nuclei since their discovery in comet 1P/Halley⁸. Volatile ices at or near the surface of a comet sublimate when exposed to sunlight during the diurnal cycle of the comet. As the gas moves outwards, it and daughter species produced by photodissociation trace spirals or arcs that can be used to infer the rotation of the comet. Cyanogen is one of the most effective gases in this respect, owing to its large fluorescence efficiency in sunlight. Its use is widespread⁹ and its connection to the rotation of comet nuclei has been verified by *in situ* missions such as EPOXI¹⁰. During our first epoch of observations (Extended Data Table 2), we identified rotating spiral arms, of which one is persistent whereas a second is visible for only part of the rotation (Fig. 1). The repetition of these features indicates a rotation period of 19.75–20.05 h during 6–9 March¹¹.

For our second epoch, we adopted a photometric technique, using variations in the brightness of the comet to measure periodicity. Although these two techniques measure different characteristics, they both identify repetitions in their respective phenomena and we assume that the associated periodicity reflects the rotation of the nucleus. We used Swift/UVOT to observe comet 41P on 7–9 May 2017 and

measured all of the light within 1,600 km of the nucleus, including molecular emissions and sunlight reflected by dust grains. The light contributed by the small nucleus was negligible during this time, indicating that variations in brightness in our aperture were dominated by the material that had recently been released from the nucleus. The photometric variations are small and slowly varying (Fig. 2, Extended Data Table 2). Although the light curve is incomplete, the unobserved parts can reasonably be inferred, resulting in a single-peaked sinusoid (the hallmark of activity being modulated by changes in illumination induced by rotation) with a period of 46–60 h. The 14-h range arises because the alignment of the overlapping segment of the phased sine curve is affected by changes in the activity of the comet with its increasing distance to the Sun. We therefore conclude that during the two months of our observations, there was a substantial change in the rotation period, with an average increase of 0.40–0.67 h per day. For the discussion presented here, we adopt 53 h, the middle of our range, as our representative period.

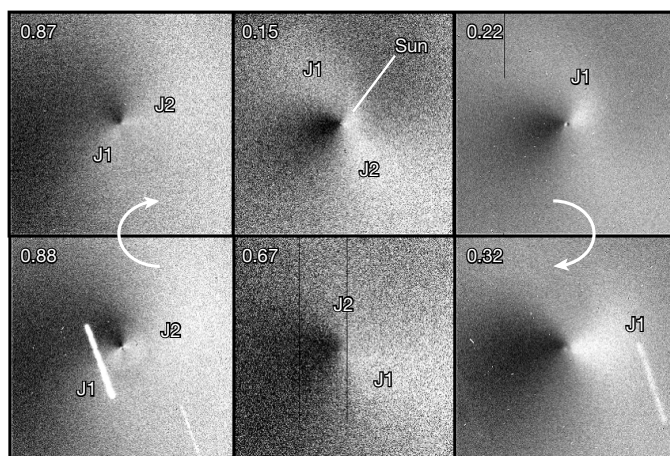


Figure 1 | Repeating cyanogen jets in the coma of comet 41P/Tuttle–Giacobini–Kresák. Sequence of DCT images showing the cyanogen coma of comet 41P, enhanced to reveal two rotating jets (labelled J1 and J2). The images, obtained on 7 and 8 March 2017, progress in a clockwise direction, as indicated by the curved arrows. Nearly identical morphologies are seen in the left two panels, which were obtained 20.1 h apart, and the sequence suggests that these two images are slightly more than one full rotation apart, leading to the derived 19.9-h period. The other panels, labelled in the upper left corner with the fraction of the period (phase) when the image was obtained, show a continuously changing morphology that precludes any periods that are sub-multiples of the 19.9-h derived value. Each panel spans approximately 20,000 km at the distance of the comet, is centred on the position of the nucleus (too small to be resolved), and is oriented with north up and east to the left. The direction to the Sun is indicated in the middle-top panel. Images were enhanced by dividing out the averaged azimuthal profile. Regions that are brighter than average at that distance from the nucleus are white and regions that are fainter are black. The white streaks are background stars.

¹Department of Astronomy, University of Maryland, College Park, Maryland 20742, USA.

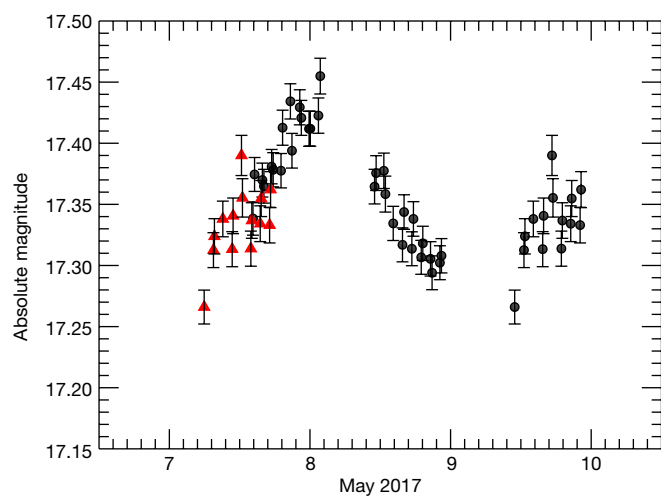


Figure 2 | Light curve of the inner coma measured by Swift/UVOT on 7–10 May 2017. The data acquired on 9.4–10 May are repeated as red triangles, phase-shifted to best match the data acquired on 7.5–8 May (black circles). Depending on the decrease in the activity of the comet with heliocentric distance (Methods), a range of periods of 46–60 h is found (Extended Data Fig. 2). The central period (53 h) is shown here. Error bars indicate 1σ stochastic uncertainties.

A cyanogen morphology similar to that seen in our DCT observations was observed on 18–27 March 2017¹², but the structure took 24 h to repeat on 19 and 21 March, and increased continuously to nearly 27 h on 26 and 27 March (Fig. 3). During the densest coverage in late March, the morphology repeated at progressively later times on subsequent nights, revealing a daily trend that is consistent with our ensemble dataset from March to May. The consistent repetition of the morphology at the end of each lengthening period over such an extended time suggests that any non-principal axis component of rotation is small. Furthermore, we cannot conceive a scenario in which non-principal axis rotation could mimic the observed continuously changing period. Therefore, we assume that the nucleus was in a state of simple rotation.

Rotation periods have been measured for scores of comets, many with extensive coverage, but 41P is only the eighth comet for which a conclusive change in period has been measured. Furthermore, the fractional change and rate of change in period for comet 41P far exceed those observed in other comets (Extended Data Table 3). Changes in rotation period depend on the size, shape, internal structure, activity and rotational state of the nucleus of the comet^{1,2,4,5}. The radius of the nucleus of comet 41P is $0.7\text{--}1.0\text{ km}$, which is less than 70%–90% of all measured radii of Jupiter-family comets¹⁴. The water production rate of comet 41P peaked around 2×10^{29} molecules per second in 2001 and 2×10^{28} molecules per second in 2006¹⁵. Our Swift observations suggest that production rates in 2017 were similar to those in 2006 (Extended Data Fig. 1). This result implies that more than 50% of the surface of the comet could be active, whereas most comets have less than 3% surface activity¹⁶. Finally, although the 20-h rotation period of comet 41P in March was long compared to most comets, the rotation period of more than 46 h that was measured in May is among the longest known¹³. It is this combination of slow rotation, high activity and a small nucleus that contribute to the rapid changes of the rotation state of comet 41P.

However, these characteristics are not unique to comet 41P. In 2010, comet 103P/Hartley 2 had an initial period of 16.5 h, a peak water production rate three times higher than that of 41P and a smaller effective radius¹⁰ of 0.57 km. Even with the more extreme combination of these characteristics, its primary rotation period changed by only 2 h in the three months around perihelion¹⁶ (Extended Data Table 3), more than an order of magnitude less than that of 41P. Therefore, other factors must also have a role in producing the net torque in comet 41P, which is much more efficient than that in any other known comet. The Deep Impact fly-by of comet 103P allowed a close examination of the activity

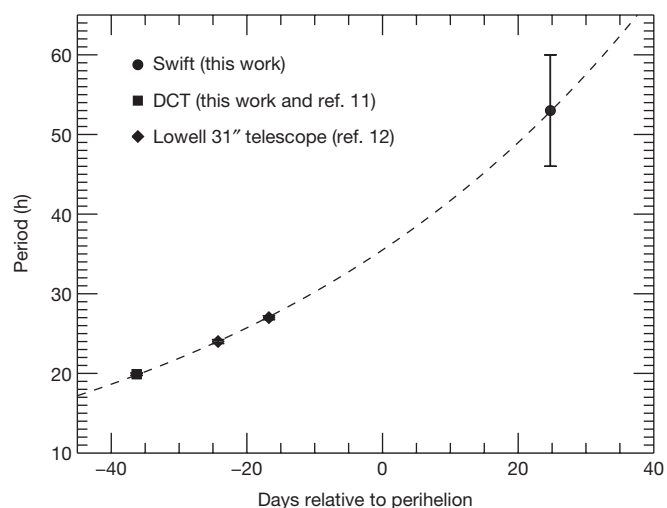


Figure 3 | Rotation periods for comet 41P measured as a function of time relative to perihelion. Perihelion occurred on 12 April 2017. The period increased at an average rate of at least 0.53 h per day over more than 60 days, an unprecedented rate of change. The different observations are indicated by symbols: Swift (circles; this work), DCT (squares; ref. 11 and this work) and results acquired¹² using the Lowell Observatory's 31" telescope (diamonds). The dashed line is a guide to the eye. Error bars indicate 1σ absolute uncertainties; the error bar on the Swift observation (circle) indicates the range of possible solutions due to the uncertainty in the change of activity as a function of heliocentric distance.

of its nucleus¹⁰, and the details that were observed enable us to explore possible differences between comets 103P and 41P. The visible jets from comet 103P are primarily along the long axis, with little moment arm for producing torques; some of the water from 103P comes from icy grains in the coma, reducing the amount of gas that contributes to torques^{18,19}; and finally, the non-principal axis rotation of 103P acts to randomize the direction of the torques, reducing their efficiency.

Using the results from the four then-available comets that exhibited changes in period, an empirical parameter X has been suggested to relate cometary activity and changes in angular momentum¹⁸. This parameter was found to be nearly constant within a range of 1–2, leading to the conclusion that net torques are nearly the same irrespective of the effective active fractions of the nucleus. From our observations of comet 41P, we compute an X parameter of more than 50, inconsistent with that conclusion. (X parameters for comets 19P/Borrelly²⁰ and 67P/Churyumov–Gerasimenko²¹ also lie well outside the 1–2 range; see Extended Data Table 3.) The deviation from this range implies that the torques, when integrated over all active areas, do not necessarily cancel out and that the physical characteristics of nuclei greatly affect the net efficiency of the torques. The effects of non-uniform activity and local topography are well illustrated by the results of the Rosetta mission to comet 67P/Churyumov–Gerasimenko, which demonstrated that the rotation period first increased, then decreased as new parts of the surface of the comet became illuminated². The active regions on the surface of comet 41P are probably oriented in such a way that its torques are highly optimized in comparison to many other comets.

We extrapolated the rotation period of the comet in time to investigate its possible past and future behaviour (Fig. 4). Our model assumes that activity levels and effective torques are constant from apparition to apparition; for example, it assumes that the orientation of the spin axis and the water production did not change substantially. It suggests that, in the near future, the rotation period could exceed 100 h. At such slow rotation rates the stabilizing gyroscope effect disappears and off-axis torques can tip the nucleus into an excited rotation state. If strong torques persist, then the rotation period of the nucleus can begin to shorten again, but with a different orientation of its rotational angular momentum vector. Such behaviour is consistent with simulations of the long-term evolution of spin states of small cometary nuclei, which

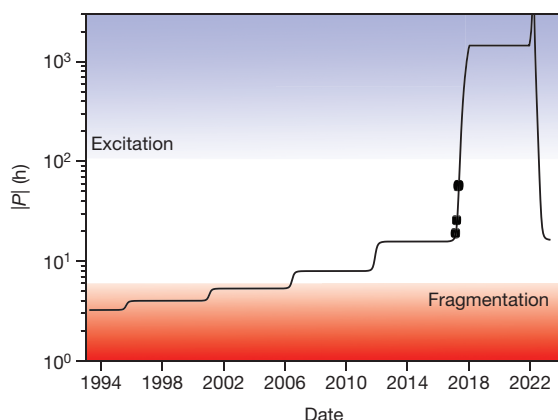


Figure 4 | Extrapolation of the rotation period of comet 41P in time.

We modelled past and future absolute values of the period (P) by extrapolating the 2017 torques to other perihelion passages (upward steps). This scenario suggests that the nucleus could evolve from rapid rotation near the fragmentation limit³ (red shading) to an excited, unstable spin state^{4,5} (blue shading) in only a few orbits. The 2017 observations are indicated by filled circles.

indicate that most comets go through a large change in their rotation period soon after their activation⁵. This large change leads to a temporary excitation of the spin state of the nucleus, and for most comets the rotation period will evolve slowly thereafter. Simulations also show that, in some cases, uniformly active surfaces can cause comets to respond unpredictably to changes in their spin state. Such comets may have inherently variable spin states, experiencing large changes in their rotation period during each perihelion passage.

Projecting back in time, comet 41P may have been near the critical fragmentation limit (with a period of around 5 h)³ in the recent past. It exhibited large outbursts in activity in 1973 and 2001^{15,22}, and these events may be related to the evolution of its spin state. The rapid rotation may have caused these outbursts via fragmentation or landslides²³; alternatively, the outbursts may have given rise to the spin evolution by exposing new active areas that generate outgassing torques.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 22 June; accepted 11 November 2017.

- Jewitt, D. Cometary rotation: an overview. *Earth Moon Planets* **79**, 35–53 (1997).
- Keller, H. U., Mottola, S., Skorov, Y. & Jorda, L. The changing rotation period of comet 67P/Churyumov–Gerasimenko controlled by its activity. *Astron. Astrophys.* **579**, L5 (2015).
- Davidsson, B. J. R. Tidal splitting and rotational breakup of solid biaxial ellipsoids. *Icarus* **149**, 375–383 (2001).
- Samarasinha, N. H., Mueller, B. E. A., Belton, M. J. S. & Jorda, L. in *Comets II* (eds Festou, M. et al.) 281–299 (Univ. Arizona Press, 2004).
- Gutiérrez, P. J., Jorda, L., Ortiz, J. L. & Rodrigo, R. Long-term simulations of the rotational state of small irregular cometary nuclei. *Astron. Astrophys.* **406**, 1123–1133 (2003).
- Gehrels, N. et al. The Swift Gamma-Ray Burst mission. *Astrophys. J.* **611**, 1005–1020 (2004).

- Farnham, T. L., Schleicher, D. G. & A'Hearn, M. F. The HB narrowband comet filters: standard stars and calibrations. *Icarus* **147**, 180–204 (2000).
- A'Hearn, M. F. et al. Cyanogen jets in comet Halley. *Nature* **324**, 649–651 (1986).
- Schleicher, D. G. & Farnham, T. L. in *Comets II* (eds Festou, M. et al.) 449–469 (Univ. Arizona Press, 2004).
- A'Hearn, M. F. et al. EPOXI at comet Hartley 2. *Science* **332**, 1396–1400 (2011).
- Farnham, T. L. et al. Comet 41P/Tuttle–Giacobini–Kresák. *CBET* **4375** (2017).
- Knight, M. M., Eisner, N., Schleicher, D. G. & Thirouin, A. Comet 41P/Tuttle–Giacobini–Kresák. *CBET* **4377** (2017).
- Lamy, P. L., Toth, I., Fernandez, Y. R. & Weaver, H. A. in *Comets II* (eds Festou, M. et al.) 223–264 (Univ. Arizona Press, 2004).
- Fernández, Y. R. et al. Thermal properties, sizes, and size distribution of Jupiter-family cometary nuclei. *Icarus* **226**, 1138–1170 (2013).
- Combi, M. SOHO SWAN derived cometary water production rates collection, urn:nasa:pds:soho:swan_derived:1.0 (ed. Feaga, L.) *NASA Planetary Data System* (2017).
- A'Hearn, M. F., Millis, R. L., Schleicher, D. G., Osip, D. J. & Birch, P. V. The ensemble properties of comets: Results from narrowband photometry of 85 comets, 1976–1992. *Icarus* **118**, 223–270 (1995).
- Knight, M. M. & Schleicher, D. G. CN morphology studies of comet 103P/Hartley 2. *Astron. J.* **141**, 183 (2011).
- Samarasinha, N. H. & Mueller, B. E. A. Relating changes in cometary rotation to activity: current status and applications to comet C/2012 S1 (ISON). *Astrophys. J.* **775**, L10 (2013).
- Belton, M. J. Cometary evolution and cryovolcanism. *Can. J. Phys.* **90**, 807–815 (2012).
- Mueller, B. E. A. & Samarasinha, N. H. Further investigation of changes in cometary rotation. In *Asteroids, Comets, Meteors 2017 meeting*, abstr. 1.e.32, <http://acm2017.uu/abstracts/Poster1.e.43.pdf> (2017).
- Mottola, S. et al. The rotation state of 67P/Churyumov–Gerasimenko from approach observations with the OSIRIS cameras on Rosetta. *Astron. Astrophys.* **569**, L2 (2014).
- Kronk, G. W. 41P/Tuttle–Giacobini–Kresák. Gary W. Kronk's *Cometography* <http://cometography.com/pcomets/041p.html> (2017).
- Steckloff, J. K., Graves, K., Hirabayashi, T., Melosh, H. J. & Richardson, J. Rotationally induced surface slope instabilities and the activation of CO₂ activity on comet 103P/Hartley 2. *Icarus* **272**, 60–69 (2016).

Acknowledgements We thank M. Siegel and the Swift team for planning the observations of 41P. This research was supported by Swift Guest Investigator Program grant 1316125. We thank A. Thirouin, C. Trujillo and N. Moskovitz for observing and/or donating telescope time to acquire images used to determine rotation periods from morphology. We thank N. Eisner and D. Schleicher for sharing their preliminary results with us. We thank N. Samarasinha for calculating the ζ parameter for 41P and 67P. This work made use of the Discovery Channel Telescope at Lowell Observatory. Lowell is a private, non-profit institution dedicated to astrophysical research and public appreciation of astronomy and operates the DCT in partnership with Boston University, the University of Maryland, the University of Toledo, Northern Arizona University and Yale University. The Large Monolithic Imager was built by Lowell Observatory using funds provided by the National Science Foundation (AST-1005313). This work also made use of NASA's Astrophysics Data System and of the JPL/Horizons ephemerides service, maintained by the JPL Solar System Dynamics group.

Author Contributions D.B. and T.L.F. designed and analysed the Swift observations. D.B., T.L.F. and M.S.P.K. planned and acquired the DCT observations. T.L.F. processed and analysed the DCT data. M.S.P.K. and D.B. modelled the change in period. All authors wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to D.B. (dennis@astro.umd.edu).

Reviewer Information *Nature* thanks B. E. A. Mueller and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Photometry. Swift/UVOT observations were obtained with the V-band filter, centred at 547 nm with a full-width at half-maximum (FWHM) of 75 nm. We measured the brightness of the coma using photometry extracted from a circular aperture centred on the nucleus with a 1,600-km (10–11-arcsec) radius at the distance of the comet. The median background flux was determined from an annulus with an inner radius of 50 arcsec and an outer radius of 100 arcsec (beyond the visible extent of the coma). We followed a standard calibration procedure²⁴ to derive the apparent magnitudes, V . These were then converted into absolute magnitudes, H , at 1 AU to account for changes in the geocentric distance Δ , heliocentric distance r_h and phase angle PA (using a phase function normalized to a phase angle of 90°, $PA(90)$)²⁵ of the comet during our observation using the relation:

$$H = V - 5\log(\Delta) - 5\log(r_h) - 2.5\log[PA/PA(90)] \quad (1)$$

The relation between the activity of the comet and its heliocentric distance, which increased from 1.099 AU to 1.108 AU during the Swift observations, is currently not well constrained. This implies that a range of scale factors A are possible for the activity-corrected brightness H' of the comet:

$$H' = H - A\log(r_h/r_0) \quad (2)$$

where r_0 is the heliocentric distance of the comet at the first Swift observation (1.099 AU). Larger scale factors imply longer rotation periods. We considered scale factors of $A = 0$ (an r_h^2 relation; see equation (1)), $A = 28$ (an early empirical fit to the current brightness trend²⁶), and an upper limit of $A = 35$ (derived from an empirical fit to the brightness trend during the apparitions of 1995 and 2001²⁶). As is shown in Extended Data Fig. 2, this results in a range of possible periods of repetition between 46 h and 60 h, with a central solution around 53 h ($A = 17$). Independent of the r_h correction, periods shorter than 46 h are not possible with our light curve (under our assumptions of simple rotation and no outburst or other unusual activity).

There are too few measurements with the DCT to construct a meaningful light curve, and the night of 8 March was not photometric (owing to Cirrus clouds); consequently, our observations focus on morphology rather than absolute measurements.

Production rates. We used Swift/UVOT images to determine water production rates following a previously outlined method²⁴. The UVW1 filter (central wavelength, 260 nm; FWHM, 70 nm) encompasses the three strong OH A–X transitions. We first created stacks containing all UVW1 images and V-band images acquired on 4–9 May 2017 using a clipped mean routine. We then removed the continuum contribution to the light measured with this filter by subtracting a weighted V-band image. There was no obvious repetitive morphology in the OH images. Fluxes in apertures with radii of 5–300 arcsec (775–46,500 km at the comet) were converted into OH column densities assuming fluorescence rates at the heliocentric velocity and distance of the comet²⁷. Production rates were derived using a vectorial model²⁸.

Active area. We derived the minimum active area corresponding to the measured water production rate using a sublimation model²⁹. We assume that every surface element has constant solar elevation—as would be the case if the spin axis were pointed at the Sun (an obliquity of 90 degrees) or if the nucleus were rotating very slowly—and is therefore in local, instantaneous equilibrium with sunlight. This maximizes the sublimation averaged over the entire surface and results in a minimum total active area. We further assumed a Bond albedo of 0.02 and 100% infrared emissivity. The modelled H₂O sublimation rate is 3.35×10^{17} molecules per cm² at 1.05 AU. Assuming a peak water production rate of 2×10^{28} molecules per second (Extended Data Fig. 1), we find an active area of at least 6 km², equivalent to an active fraction of 50%–97% of a spherical nucleus with a radius of 0.7–1 km.

Modelling the change in rotation period. To extrapolate the rotation period of comet 41P to past and future apparitions we used the relation between the rate of change of the angular velocity $d\omega/dt$, the water mass loss rate Q and the radius of the nucleus R (ref. 18):

$$\frac{d\omega}{dt} = C \frac{Q(t)}{R^4} \quad (3)$$

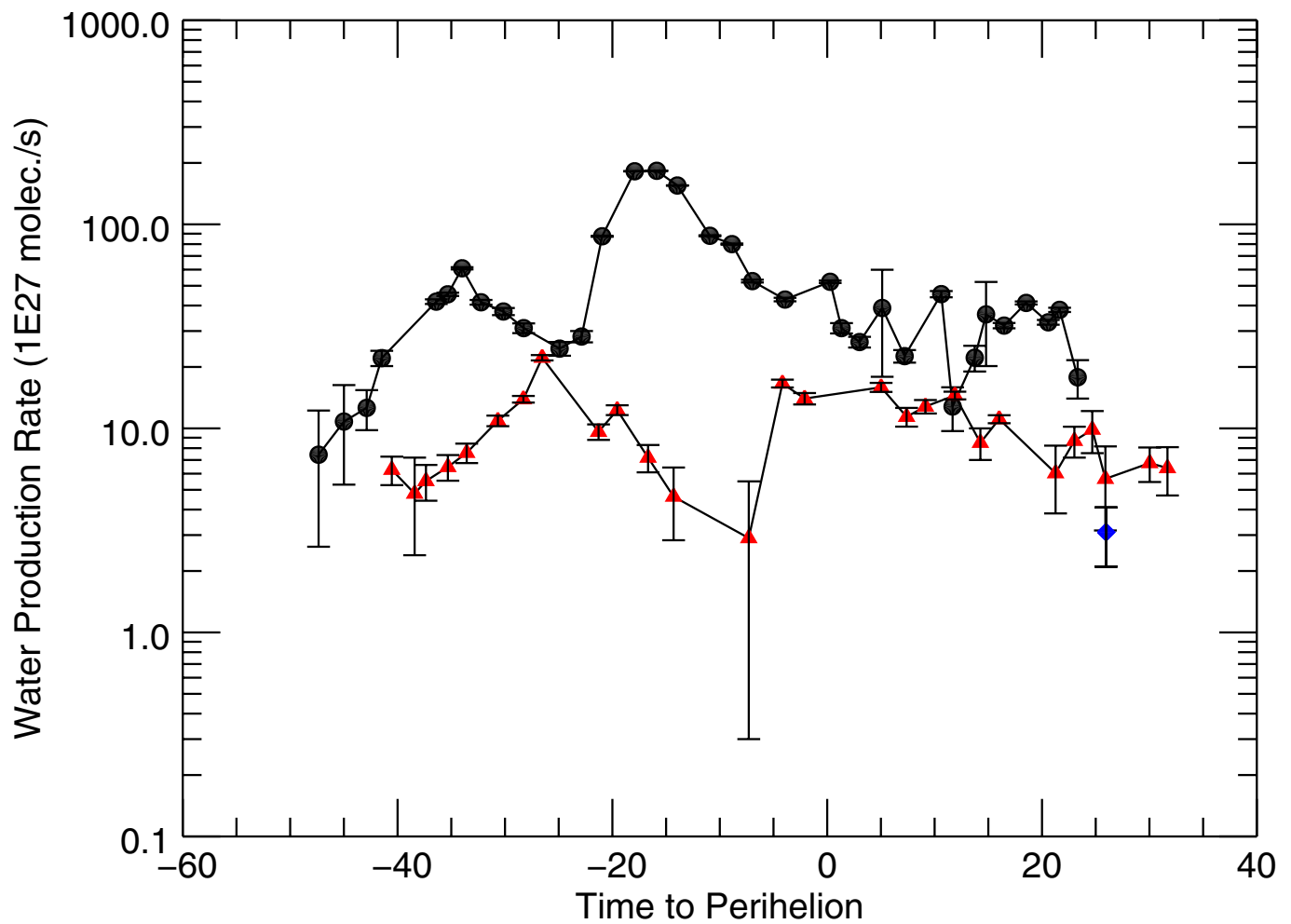
We assumed a nucleus with a radius of 0.7 km and used our measurements of the production rate and the average change of rotation period during the current apparition to determine the constant C empirically. To estimate the orbital gas mass loss, we fitted the empirical relation between the brightness of the comet and its heliocentric distance ($Q \propto r_h^{-4.8}$) to the SOHO measurements of water production rates during the 2006 apparition¹⁵. We assumed abundances of 10% for both CO and CO₂ relative to water, and that activity beyond 3 AU is negligible. When the nucleus reaches a rotation frequency of 0, the period is infinite, hence the growth off the top of Fig. 4. At this point in the model, the rotation reverses (rotational pole flip) and the period decreases from infinity. However, in reality the rotation will become excited, the illumination on the surface will change and the torques should also change.

Integrating the gas production rates from 3 AU before to 3 AU after perihelion results in a mass loss rate of 6×10^9 kg in volatiles per orbit, or about 1% of the mass of the nucleus for a density of 500 kg m⁻³.

Rotation evolution models⁵ assume a certain initial spin state and evolution is modelled for 10–100 orbits. Comet 41P has orbited the Sun approximately 30 times since its discovery in 1858. After considering several scenarios, hyperbolic evolution after 10–30 orbits was concluded previously, with the spin states of the comets evolving continuously throughout the simulations⁵. However, these models⁵ did not explore the full parameter space⁴ and we are hesitant to imply a more quantitative interpretation of them.

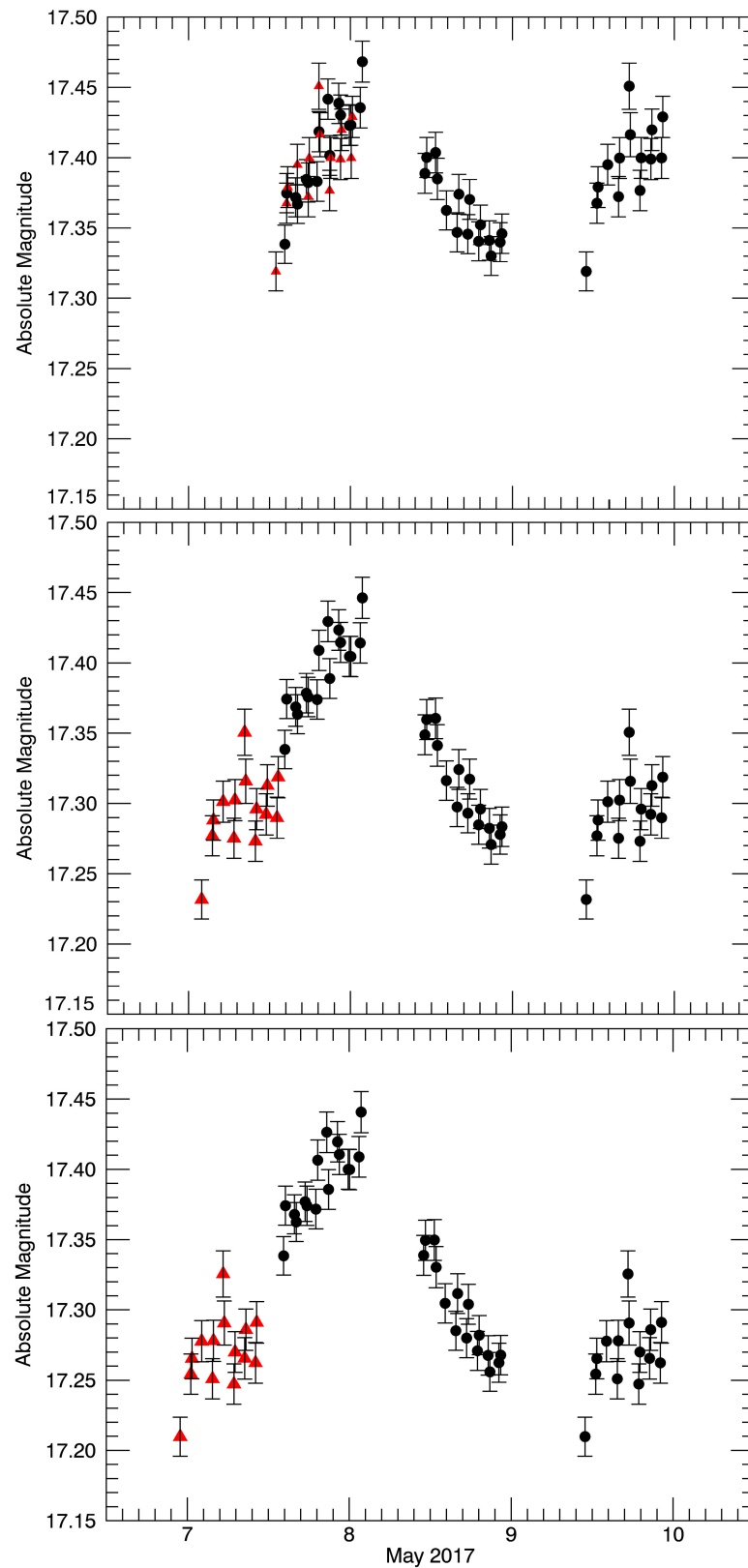
Data availability. All Swift/UVOT data are available from the Barbara A. Mikulski Archive for Space Telescopes (<https://archive.stsci.edu>) and from the Swift Archive Portal (http://www.swift.ac.uk/swift_portal/) under programme ID 1316125. The photometric measurements are provided as Source Data for Fig. 2. Other data that support the findings of this study are available from the corresponding author on reasonable request.

24. Bodewits, D. *et al.* The evolving activity of the dynamically young comet C/2009 P1 (Garradd). *Astrophys. J.* **786**, 48 (2014).
25. Schleicher, D. Composite dust phase function for comets. *Lowell Observatory* <http://asteroid.lowell.edu/comet/dustphase.html> (2010).
26. Yoshida, S. 41P/Tuttle–Giacobini–Kresák. *Seichi Yoshida's Home Page* <http://www.aerith.net/comet/catalog/0041P/2017.html> (2017).
27. Schleicher, D. G. & A'Hearn, M. F. The fluorescence of cometary OH. *Astrophys. J.* **331**, 1058–1077 (1988).
28. Festou, M. C. The density distribution of neutral compounds in cometary atmospheres. I. Models and equations. *Astron. Astrophys.* **95**, 69–79 (1981).
29. Cowan, J. J. & A'Hearn, M. F. Vaporization of comet nuclei: light curves and life times. *Moon Planets* **21**, 155–171 (1979).
30. A'Hearn, M. F. *et al.* Deep Impact: excavating comet Tempel 1. *Science* **310**, 258–264 (2005).
31. Knight, M. M., Schleicher, D. G., Farnham, T. L., Schwieterman, E. W. & Christensen, S. R. A quarter-century of observations of comet 10P Tempel 2 at Lowell Observatory: continued spin-down, coma morphology, production rates, and numerical modeling. *Astrophys. J.* **144**, 153 (2012).
32. Mueller, B. E. A. & Ferrin, I. Change in the rotational period of comet P/Tempel 2 between the 1988 and 1994 apparitions. *Icarus* **123**, 463–477 (1996).
33. Mueller, B. E. A., Samarasinha, N. H., Rauer, H. & Helbert, J. Determination of a precise rotation period for the Deep Space 1 target, comet 19P/Borely. *Icarus* **209**, 745–752 (2010).
34. Sierks, H. *et al.* On the nucleus structure and activity of comet 67P/Churyumov–Gerasimenko. *Science* **347**, aaa1044 (2015).
35. ESA Flight Dynamics Team. Comet rotation period. *European Space Agency* <http://sci.esa.int/rosetta/58367-comet-rotation-period/> (2017).
36. Schleicher, D. G., Millis, R. L. & Osip, D. J. Comet Levy (1990c): ground-based photometric results. *Icarus* **94**, 511–523 (1991).
37. Feldman, P. D., Budzien, S. A., Festou, M. C., A'Hearn, M. F. & Tozzi, G. P. Ultraviolet and visible variability of the coma of comet Levy (1990c). *Icarus* **95**, 65–72 (1992).



Extended Data Figure 1 | Water production rates of comet 41P in 2001, 2006 and 2017. Production rates were derived from hydrogen Lyman- α emission observed by the SWAN instrument on board the SOHO spacecraft¹⁵ in 2001 (black circles) and 2006 (red triangles). For the SWAN data, 1σ stochastic errors are shown; systematic uncertainties are at the 30% level¹⁵. We used Swift/UVOT observations of hydroxyl (OH)

emission to determine the water production rate in 2017 (blue diamond). For the Swift data, the error bars represent the systematic uncertainty. The comet had two 4-mag outbursts in optical wavelengths just before its perihelion in 2001²²; these are evident as peaks at approximately 35 and 15 days before perihelion.



Extended Data Figure 2 | Rotation periods for different activity models.

Absolute magnitudes based on Swift/UVOT photometry (black circles) are corrected for different relationships (characterized by A) between the activity of the comet and its distance to the Sun (see Methods). An

increase in A corresponds to an increase in the rotation period that is needed to phase the overlapping sine curve segment (red triangles). Top, $A = 0$, period = 46 h; middle, $A = 28$, period = 57 h; bottom, $A = 35$, period = 60 h. Error bars indicate 1σ stochastic uncertainties.

Extended Data Table 1 | Summary of measured rotation periods

Telescope	Dates	Δ	r_h	Rotation Period	References
	(UTC)	(au)	(au)	(hours)	
DCT/LMI	Mar. 6 – Mar. 9, 2017	0.24 – 0.18	1.22 - 1.16	19.9 ± 0.15	This work; 11
Lowell 31"	Mar. 18 – 27, 2017	0.16 - 0.14	1.1 - 1.06	$24 - 27 \pm 0.25$	(12)
Swift/UVOT	May 6 – 9, 2017	0.21	1.1	46 – 60	This work

Δ is the geocentric distance of the comet; r_h denotes its heliocentric distance. Data are from this work and refs 11 and 12.

Extended Data Table 2 | Observing log of Lowell Observatory's DCT

Date (UTC)	Midtime (UTC)	r_h (au)	Δ (au)	Phase (deg.)	Rot. Phase	Observers
Mar. 6, 2017	2:38	1.16	0.20	27.5	0.32	Thirouin/Moskovitz
Mar. 7, 2017	5:27	1.16	0.19	29.0	0.67	Farnham/Kelley/Bodewits
Mar. 7, 2017	9:33	1.15	0.19	29.3	0.87	Farnham/Kelley/Bodewits
Mar. 8, 2017	3:41	1.15	0.19	30.3	0.78	Farnham/Kelley/Bodewits
Mar. 8, 2017	5:40	1.15	0.19	30.4	0.88	Farnham/Kelley/Bodewits
Mar. 8, 2017	8:23	1.15	0.19	30.6	0.02	Farnham/Kelley/Bodewits
Mar. 8, 2017	11:04	1.15	0.19	30.7	0.15	Farnham/Kelley/Bodewits
Mar. 9, 2017	8:15	1.14	0.18	32.0	0.22	Thirouin/Moskovitz

Δ is the geocentric distance of the comet; r_h denotes its heliocentric distance.

Extended Data Table 3 | Characteristics of other comets for which a change in rotation period has been measured

Name	P (hr)	ΔP (hr)	Interval* (yr)	ΔP /orbit (hr/orbit)	Effective Radius (km)	Active Fraction (%)	X	References
41P/T-G-K	20	>26	0.2	>26	<1	>50	78	This paper; 13
2P/Encke	11	-0.072	3.3	-0.072	2.4	1.0	1.0	13, 16, 18
9P/Tempel 1	41	-0.233	5.52	-0.233	2.8	6.5	1.1	16, 18, 30
10P/Tempel 2	9	0.018	22.5	0.045	5.3	0.8	1.9	13, 16, 31, 32
19P/Borelly	29	>0.667	13	0.33	2.4	9.1	5.1	16, 20, 31, 33
67P/C-G (pre) **)	12	0.027	1.21	0.027	1.65	2.0	0.4	2, 21, 34, 35
67P/C-G (post) **)		-0.375	1.33	-0.375			6.1	2, 21, 34, 35
103P/Hartley 2	17	2	0.25	2	0.57	>100	1.5	13, 17, 18, 19
Levy (1990c)	19	-1.3	0.058	-1.3	-	-	-	36, 37

Data are from this work and refs 2, 13, 16–21 and 30–37.

*Interval between measurements of rotation period, which may not reflect the time it took to change. In some instances, changes in period have been observed on multiple orbits.

**For 67P/Churyumov–Gerasimenko (67P/C-G), characteristics before and after perihelion are given separately.

Bright triplet excitons in caesium lead halide perovskites

Michael A. Becker^{1,2*}, Roman Vaxenburg^{3*}, Georgian Nedelcu^{4,5}, Peter C. Sercel⁶, Andrew Shabaev³, Michael J. Mehl⁷, John G. Michopoulos⁸, Samuel G. Lambrakos⁸, Noam Bernstein⁸, John L. Lyons⁸, Thilo Stöferle¹, Rainer F. Mahrt¹, Maksym V. Kovalenko^{4,5}, David J. Norris², Gabriele Rainò^{1,4} & Alexander L. Efros⁸

Nanostructured semiconductors emit light from electronic states known as excitons¹. For organic materials, Hund's rules² state that the lowest-energy exciton is a poorly emitting triplet state. For inorganic semiconductors, similar rules³ predict an analogue of this triplet state known as the 'dark exciton'⁴. Because dark excitons release photons slowly, hindering emission from inorganic nanostructures, materials that disobey these rules have been sought. However, despite considerable experimental and theoretical efforts, no inorganic semiconductors have been identified in which the lowest exciton is bright. Here we show that the lowest exciton in caesium lead halide perovskites (CsPbX₃, with X = Cl, Br or I) involves a highly emissive triplet state. We first use an effective-mass model and group theory to demonstrate the possibility of such a state existing, which can occur when the strong spin–orbit coupling in the conduction band of a perovskite is combined with the Rashba effect^{5–10}. We then apply our model to CsPbX₃ nanocrystals¹¹, and measure size- and composition-dependent fluorescence at the single-nanocrystal level. The bright triplet character of the lowest exciton explains the anomalous photon-emission rates of these materials, which emit about 20 and 1,000 times faster¹² than any other semiconductor nanocrystal at room^{13–16} and cryogenic⁴ temperatures, respectively. The existence of this bright triplet exciton is further confirmed by analysis of the fine structure in low-temperature fluorescence spectra. For semiconductor nanocrystals, which are already used in lighting¹⁷, lasers¹⁸ and displays¹⁹, these excitons could lead to materials with brighter emission. More generally, our results provide criteria for identifying other semiconductors that exhibit bright excitons, with potential implications for optoelectronic devices.

An exciton involves an electron in the conduction band that is bound Coulombically to a hole in the valence band. Its energy depends in part on the spin configuration of these two charge carriers. In organic semiconductors, the lowest-energy exciton is a triplet state in which these two carriers have parallel spins. For the electron and hole to recombine and emit light, one spin must flip simultaneously with the release of the photon to satisfy the Pauli exclusion principle. Because this coordinated process is unlikely, triplet excitons are poorly emitting.

In addition to spin, the exciton energy depends on the atomic orbitals that constitute the conduction and valence bands. In many inorganic semiconductors, the orbital motion and spin of the carriers are strongly coupled. Spin is no longer conserved, and the total angular momentum of the electron and hole (J_e and J_h) must be considered. Further, the exchange interaction mixes these momenta so that only the total exciton momentum $J = J_e + J_h$ is conserved. Owing to these and other effects, each exciton state is split into several energy sublevels, known

as fine structure. Studies on various materials have found that the lowest-energy sublevel is 'dark', meaning that optical transitions to the ground state are dipole-forbidden. Emission, if it occurs, is very slow. For example, in CdSe, recombination of the lowest exciton requires a change of two units of angular momentum⁴. Because the photon carries one unit, light cannot be emitted unless another unit is dissipated simultaneously, another unlikely process. The lowest excitons in all known inorganic semiconductors behave similarly, leading to the common belief that such states must be dark.

We show that this belief is incorrect by examining CsPbX₃ (X = Cl, Br or I) perovskites. Crystals of these perovskites comprise corner-sharing PbX₆ octahedra with Cs⁺ ions filling the voids in between (Fig. 1a). We first approximate the lattice as cubic and calculate band structures (Methods) for CsPbBr₃ (Fig. 1b), CsPbCl₃ and CsPbI₃ (Extended Data Fig. 1). The bandgap occurs at the R point in the Brillouin zone, near which the valence and conduction bands are well described within the effective-mass model (see Supplementary Table 1). The top of the valence band arises from a mixture of Pb 6s and Br 4p atomic orbitals, with an overall *s* symmetry^{20,21}. Thus, including spin, the hole can occupy one of two *s*-like Bloch states with $J_h = 1/2$: $|\uparrow\rangle_h = |S\rangle|\uparrow\rangle$ or $|\downarrow\rangle_h = |S\rangle|\downarrow\rangle$, using standard notation²². The conduction band consists of Pb 6p orbitals, leading to three possible orthogonal spatial components for the Bloch function^{20,21}: $|X\rangle$, $|Y\rangle$ or $|Z\rangle$. Because of strong spin–orbit coupling, these components are mixed with spin to obtain a doubly degenerate $J_e = 1/2$ state for the electron at the bottom of the conduction band:

$$\begin{aligned} |\uparrow\rangle_e &= -\frac{1}{\sqrt{3}}[(|X\rangle + i|Y\rangle)|\downarrow\rangle + |Z\rangle|\uparrow\rangle] \\ |\downarrow\rangle_e &= \frac{1}{\sqrt{3}}[|Z\rangle|\downarrow\rangle - (|X\rangle - i|Y\rangle)|\uparrow\rangle] \end{aligned} \quad (1)$$

When the momentum of the electron and hole states are then combined, the exciton splits as a result of electron–hole exchange into a $J = 0$ singlet state

$$|\Psi_{0,0}\rangle = \frac{1}{\sqrt{2}}[|\downarrow\rangle_e|\uparrow\rangle_h - |\uparrow\rangle_e|\downarrow\rangle_h] \quad (2)$$

and a threefold degenerate $J = 1$ triplet state

$$\begin{aligned} |\Psi_{1,-1}\rangle &= |\downarrow\rangle_e|\downarrow\rangle_h \\ |\Psi_{1,0}\rangle &= \frac{1}{\sqrt{2}}[|\downarrow\rangle_e|\uparrow\rangle_h + |\uparrow\rangle_e|\downarrow\rangle_h] \\ |\Psi_{1,+1}\rangle &= |\uparrow\rangle_e|\uparrow\rangle_h \end{aligned} \quad (3)$$

¹IBM Research – Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland. ²Optical Materials Engineering Laboratory, Department of Mechanical and Process Engineering, ETH Zurich, 8092 Zurich, Switzerland. ³Computational Materials Science Center, George Mason University, Fairfax, Virginia 22030, USA. ⁴Institute of Inorganic Chemistry, Department of Chemistry and Applied Biosciences, ETH Zurich, 8093 Zurich, Switzerland. ⁵Laboratory for Thin Films and Photovoltaics, Empa, Swiss Federal Laboratories for Materials Science and Technology, CH-8600 Dübendorf, Switzerland. ⁶T. J. Watson Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA. ⁷US Naval Academy, Annapolis, Maryland 21402, USA. ⁸Center for Computational Materials Science, Naval Research Laboratory, Washington DC 20375, USA.

*These authors contributed equally to this work.

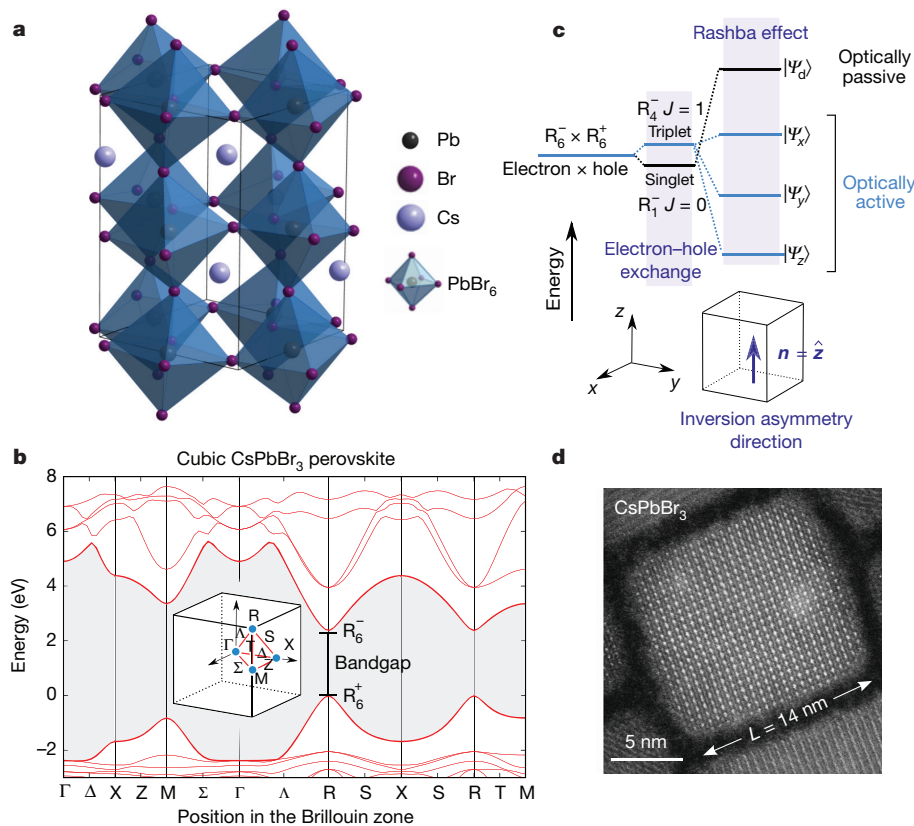


Figure 1 | Crystal and electronic structure for CsPbBr₃ perovskite. **a**, Orthorhombic crystal structure of CsPbBr₃ (*Pnma* space group, unit cell shown as a frame), which differs from the idealized cubic perovskite by an octahedral tilting. **b**, Calculated band structure of cubic CsPbBr₃ perovskite. The inset shows the first Brillouin zone of the cubic crystal lattice. The electronic bandgap is indicated in the band structure at the R point. The valence (conduction) band maximum (minimum) has R_6^+ (R_6^-) symmetry. **c**, The expected fine structure of the band-edge exciton considering short-range electron–hole exchange (middle) and then including the Rashba effect (right) under orthorhombic symmetry. The latter splits the exciton into three bright states with transition dipoles oriented along the orthorhombic symmetry axes (labelled *x*, *y* and *z*) and a higher-energy dark state (labelled ‘d’). The energetic order of the three lowest sublevels is determined by the orthorhombic distortion. The orthorhombic unit cell (bottom) and the resulting sublevel order is shown for CsPbBr₃. **d**, Transmission electron micrograph of an individual CsPbBr₃ nanocrystal with an edge length of $L = 14$ nm.

where each $|\Psi_{J_z}\rangle$ is labelled with J_z , the *z* projection of *J*. The probability of light emission due to electron–hole recombination from these excitons can then be calculated (Supplementary Information section 1). We find a probability of zero for $|\Psi_{0,0}\rangle$ and of non-zero for $|\Psi_{1,J_z=0,\pm1}\rangle$, indicating a dark singlet and a bright triplet.

These selection rules are confirmed by group theory. At the R point, the band-edge electron and hole states transform as irreducible representations R_6^- and R_6^+ , respectively (with the superscript denoting parity)²³. Exchange then splits the exciton into a dark singlet (R_1^-) and a bright triplet (R_4^-); see Supplementary Information section 2 and Supplementary Table 3.

Detailed calculations (see Supplementary Information section 1) reveal the energetic order of these levels. If only short-range exchange is included, then the singlet lies below the triplet (Fig. 1c). However, CsPbX₃ perovskites should also exhibit a large Rashba effect⁵. This occurs in semiconductors with strong spin–orbit coupling and an inversion asymmetry. For the closely related hybrid organic–inorganic perovskites, the impact of this effect on photovoltaic and spintronic devices has been discussed extensively^{6–9}. Although the cause of the inversion asymmetry (cation positional instabilities²⁴ or surface effects⁹) remains unknown, the Rashba effect should alter the fine structure. Indeed, the bright triplet exciton can be lowered below the dark singlet exciton.

To examine this possibility, we studied colloidal nanocrystals of CsPbX₃ (Methods). Compared to bulk crystals, nanocrystals enable the additional effect of system size to be investigated. Such particles are roughly cube-shaped with edge lengths of $L = 8$ –15 nm (Fig. 1d). Before these nanocrystals were introduced¹¹, all technologically relevant semiconductor nanocrystals exhibited slow, sub-microsecond radiative lifetimes at cryogenic temperatures, owing to the lowest exciton being dark⁴. By contrast, CsPbX₃ nanocrystals emit about 1,000 times faster (with sub-nanosecond lifetimes)¹². In Fig. 2a we show photoluminescence decays for individual CsPbI₃, CsPbBr₃ and CsPbBr₂Cl nanocrystals at cryogenic temperatures. The decay times are 0.85 ns, 0.38 ns and 0.18 ns, respectively, decreasing with increasing emission energy. The photoluminescence quantum yield for the fastest

of these samples, the CsPbBr₂Cl nanocrystals ($L = 14 \pm 1$ nm; throughout we quote the mean value and standard deviation from several measurements), was measured to be near unity ($88\% \pm 14\%$) at 5 K (Extended Data Fig. 2), which indicates that these decay times can be related directly to radiative lifetimes. In Fig. 2b we present a larger set of decay times (squares) for individual CsPbI₃, CsPbBr₃ and CsPbBr₂Cl nanocrystals. All are much shorter than those reported for CdSe, CdS, CdTe, InAs, InSb, InP, PbSe, PbS and PbTe nanocrystals^{13–16}, consistent with the lowest exciton being the bright triplet.

However, fast decays could also indicate emission from trions (charged excitons). Trions are optically active, but suffer from rapid non-radiative Auger recombination. They should therefore exhibit quicker but weaker decays than excitons. In our single-nanocrystal experiments discussed above, trion contributions are reduced by spectral filtering (Extended Data Fig. 3). However, to test the role of trions explicitly, we analysed the photon stream from individual nanocrystals without filtering (Fig. 2c, d, left). The correlation of emission intensity with lifetime allows the strong exciton and weak trion contributions to be separated (Fig. 2c, d, right)²⁵. We confirm fast exciton lifetimes for CsPbI₃ and CsPbBr₃ nanocrystals of 1.2 ns and 0.4 ns, respectively, consistent with ensemble measurements (Extended Data Fig. 4).

To compare with theory, we calculated radiative lifetimes for perovskite nanocrystals within the effective-mass model. In addition to the wavefunctions in equations (2) and (3), exciton confinement within the nanocrystal must be included via envelope functions for the electron and hole. If CsPbX₃ nanocrystals were spherical, excitonic lifetimes could be calculated using previously described methods (Supplementary Information section 3). However, for cubes, the electric field of a photon not only changes across the boundary of the nanocrystal, owing to dielectric screening (as in spherical nanocrystals), but also becomes inhomogeneous (Fig. 2e, Extended Data Figs 5, 6). We included this inhomogeneity, along with the Rashba effect and the orthorhombic lattice distortion in CsPbX₃ nanocrystals in our calculations²⁶. For simplicity, we assumed the nanocrystals were cube-shaped. Only when the Rashba effect was included could

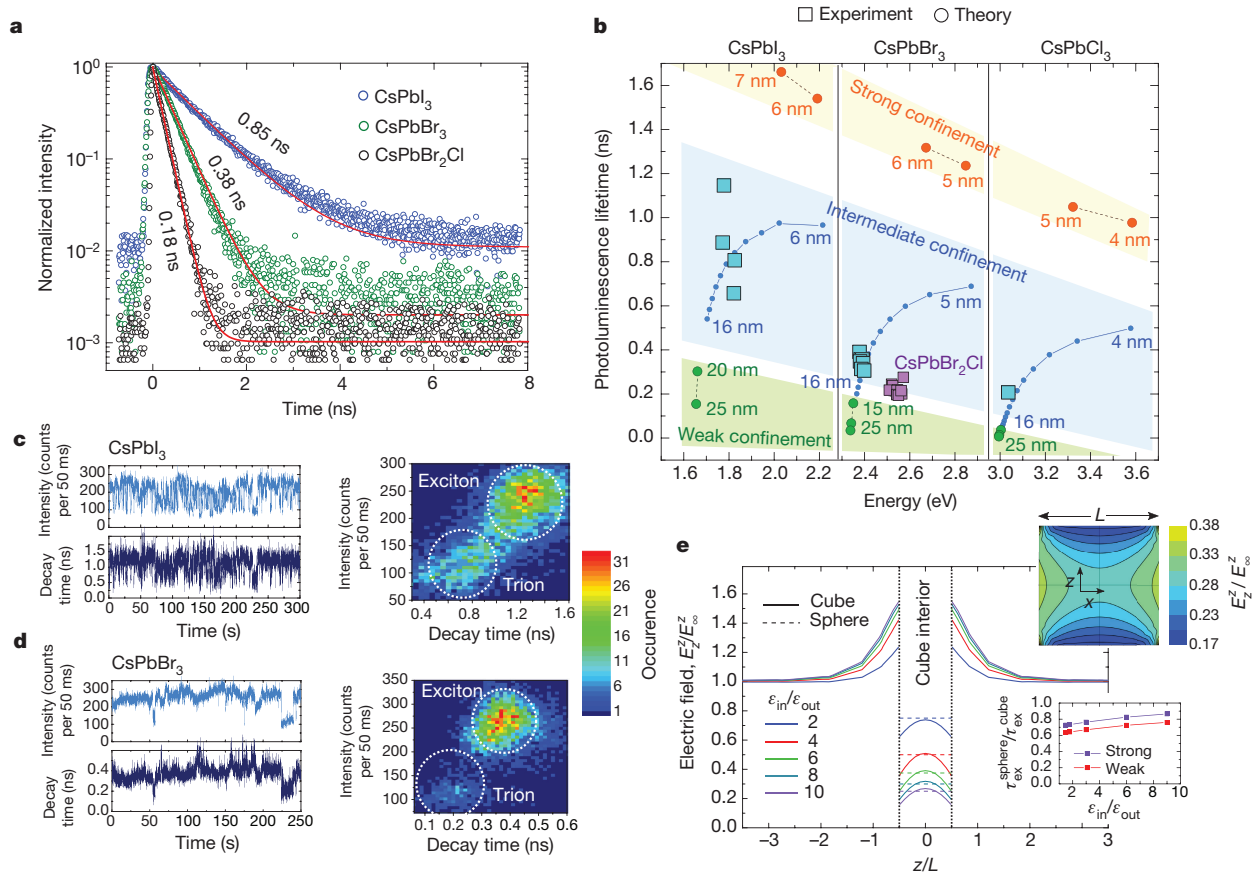


Figure 2 | Characterization of fast radiative lifetimes in CsPbX₃ nanocrystals. **a**, Photoluminescence decays (open circles) measured from single CsPbI₃ ($L = 14$ nm), CsPbBr₃ ($L = 11$ nm) and CsPbBr₂Cl ($L = 14$ nm) perovskite nanocrystals. By fitting the data with an exponential decay function (red lines), radiative decays of 0.85 ns, 0.38 ns and 0.18 ns were obtained for CsPbI₃, CsPbBr₃ and CsPbBr₂Cl perovskite nanocrystals, respectively. **b**, Calculated radiative lifetimes of the bright triplet exciton versus transition energy for CsPbX₃ nanocrystals with $X = \text{Cl, Br or I}$. The theoretical results (circles) are divided into three size regimes (labels on individual points give the edge lengths, L , of the cube-shaped nanocrystals): strong (orange), intermediate (blue) and weak (green) exciton confinement. These values are compared with measured photoluminescence decays from individual perovskite nanocrystals (squares; sizes of the CsPbI₃, CsPbBr₃ and CsPbBr₂Cl crystals as in **a**). A data point for an ensemble of CsPbCl₃ nanocrystals ($L = 10$ nm) is also shown. Measured values are consistent with calculations in the intermediate confinement regime, which include electron–hole

a self-consistent model for CsPbX₃ nanocrystals be obtained, as we describe below.

The Rashba coefficient was estimated from low-temperature photoluminescence spectra (see below). If the effective Rashba field is parallel to one of the orthorhombic symmetry axes of the nanocrystal (see Supplementary Information section 1 for details and other cases), then the bright triplet exciton ($J = 1$) is split into three non-degenerate sublevels

$$\begin{aligned} |\Psi_x\rangle &= \frac{1}{\sqrt{2}} [|\uparrow\rangle_e |\uparrow\rangle_h - |\downarrow\rangle_e |\downarrow\rangle_h] \\ |\Psi_z\rangle &= |\Psi_{1,0}\rangle \\ |\Psi_y\rangle &= \frac{1}{\sqrt{2}} [|\uparrow\rangle_e |\uparrow\rangle_h + |\downarrow\rangle_e |\downarrow\rangle_h] \end{aligned} \quad (4)$$

which lie below the dark singlet (Fig. 1c). The triplet states represent three linear dipoles polarized along the orthorhombic symmetry axes (x, y, z). Transitions from these three sublevels have the same oscillator

correlations. **c, d**, Detected photon counts (left panels) versus time from individual CsPbI₃ (**c**) and CsPbBr₃ (**d**) nanocrystals (sizes as in **a**). Traces show ‘A-type’ blinking from the nanocrystals²⁵. These data can be analysed to separate contributions to the photoluminescence decay from exciton and trion emission (right panels). The targeted temperature in all experiments was 5 K, but may be higher (10–20 K; see Fig. 3). **e**, Calculated distribution of the z component of the electric field normalized to the applied field (along the z direction) at infinite distance, E_z^z/E_∞^z . This quantity is plotted versus position z across the centre line of spherical (dashed lines) or cube-shaped (solid lines) nanocrystals for various ratios of the dielectric constant inside (ϵ_{in}) and outside (ϵ_{out}) the nanocrystals (see legend). The field inside the nanocrystal is essentially always lower for the cube than for the sphere. Upper inset, calculated two-dimensional distribution of E_z^z/E_∞^z inside a cube-shaped nanocrystals plotted on the x – z mid-plane for $\epsilon_{\text{in}}/\epsilon_{\text{out}} = 6$. Lower inset, calculated ratio $\tau_{\text{ex}}^{\text{sphere}}/\tau_{\text{ex}}^{\text{cube}}$ of radiative decay times for spherical and cubical nanocrystals with the same volume versus $\epsilon_{\text{in}}/\epsilon_{\text{out}}$ for strong (blue) and weak (red) confinement.

strength. Moreover, in cube-shaped nanocrystals, these states still emit as linear dipoles despite the inhomogeneous field (Supplementary Information sections 1 and 3).

The radiative lifetime of the triplet exciton τ_{ex} is evaluated from

$$\frac{1}{\tau_{\text{ex}}} = \frac{4\omega n E_p}{9 \times 137 m_0 c^2} I_{\parallel}^2 \quad (5)$$

with ω the angular transition frequency, n the refractive index of the surrounding medium, m_0 the free-electron mass, c the speed of light, $E_p = 2P^2/m_0$ the Kane energy (Extended Data Fig. 7) and P the Kane parameter²². I_{\parallel} is an overlap integral that includes the envelope functions of the electron and hole and the field-averaged transition-dipole moment (Supplementary Information section 3).

In Fig. 2b we present the τ_{ex} calculated for CsPbX₃ nanocrystals (circles). The results can be divided into three regimes, depending on the size of the nanocrystal. When the nanocrystal is smaller than the Bohr radius of the exciton a_B (strong exciton confinement,

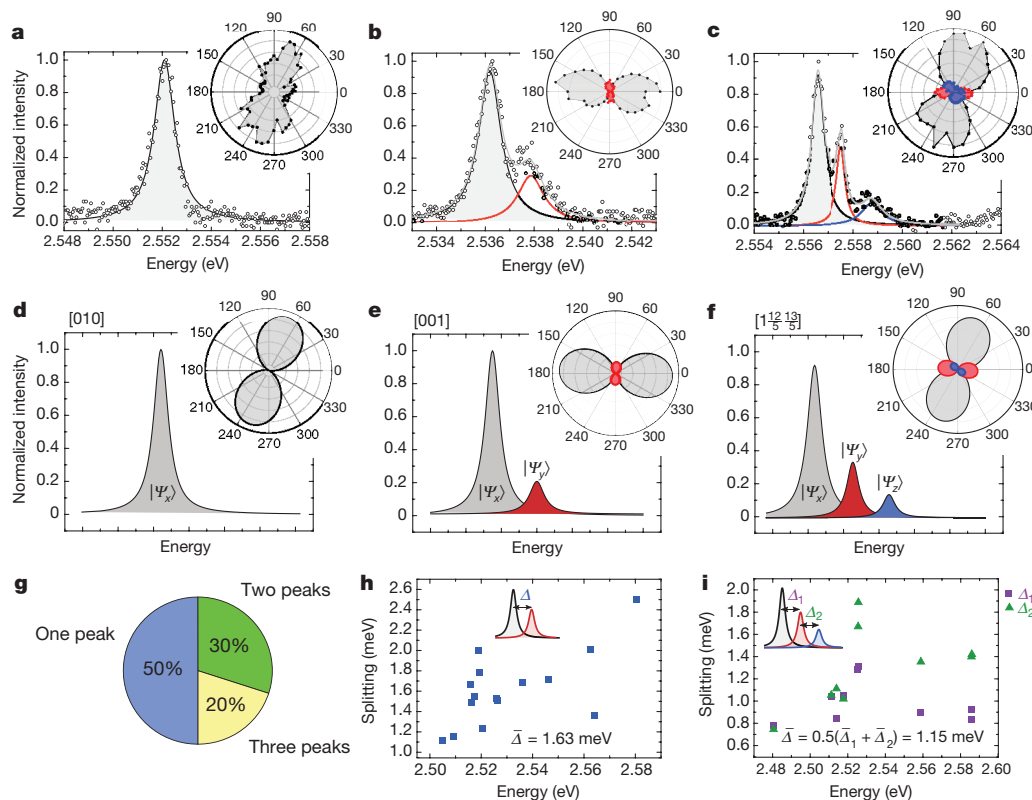


Figure 3 | Fine structure of the bright triplet exciton for CsPbBr₂Cl nanocrystals. **a–c**, Photoluminescence spectra (points) of individual nanocrystals ($L = 14 \pm 1$ nm) that exhibit a single peak (**a**), two peaks (**b**) and three (**c**) peaks. Single- and multi-Lorentzian-function fits are displayed as solid lines (grey lines are the cumulative fits). The targeted temperature was 5 K; however, the quantitative fits of the relative peak intensities, based on a Boltzmann distribution (Supplementary Information section 4), required higher temperatures (10–20 K). This may indicate a warmer sample temperature due to imperfect thermal contact and/or laser heating. Alternatively, deviations from a Boltzmann distribution may be present. The insets show the polarization of each of the spectral features. For the spectra, a linear polarizer was placed in the

detection path. The angle of this polarizer was adjusted so that the relative intensity of the features in the spectra matched the polarization dependence in the insets. **d–f**, Simulated spectra and polarizations for nanocrystal orientations that match the experimental results in **a–c**; see Supplementary Information section 4 for details. Each panel lists the required observation direction relative to the orthorhombic unit-cell axes. **g**, Experimental statistics for the observation of single-, two- and three-peak spectra from individual nanocrystals with $L = 7.5$ –14 nm (51 spectra with 35 splittings in total). **h**, **i**, Experimental fine-structure splitting measured for the two- (**h**) and three-peak (**i**) spectra. The average splitting $\bar{\Delta}$ in each case is provided.

orange circles), the predicted radiative lifetime decreases from 2 ns to 1 ns with increasing emission energy. For large nanocrystals in the opposite limit (weak exciton confinement, green circles), the lifetime should be even shorter because weakly confined excitons exhibit larger oscillator strengths²⁷. In this size regime ($L \approx 15$ –25 nm), the calculated lifetimes decrease below 100 ps for CsPbBr₃ and CsPbCl₃ nanocubes. The lifetime would be decreased further in spheres of the same volume (Fig. 2e, lower inset).

The measured photoluminescence decays in Fig. 2b (squares) lie between those predicted for strong and weak confinement. Because the size of the nanocrystals and a_B are comparable, the electron and hole motion is correlated. If this effect is taken into account (intermediate exciton confinement, blue circles), then calculations for $L \approx 4$ –16 nm (Supplementary Information section 3, Extended Data Fig. 8) agree well with the experiment.

The order of exciton levels used above depends on the values and relative signs of the Rashba coefficients for the electron and hole. If they have the same sign, then the angular-momentum texture (that is, how the orientation of the angular momentum varies with wavevector) exhibits the same helicity at the valence-band maximum and conduction-band minimum⁷. Optical transitions between these bands are allowed when the helicity is preserved (owing to their s and p symmetry, respectively.) Thus, for this case, the lowest exciton sublevel should be bright; see Supplementary Information section 1.E for details. A similar situation exists in transition-metal dichalcogenide monolayers²⁸.

We estimated the values of the Rashba coefficients from photoluminescence spectra of individual nanocrystals, which reveal the fine structure directly. Our nanocrystals exhibit one, two or three peaks, all with near-linear polarization (Fig. 3a–c, Extended Data Figs 9, 10). This is consistent with the three non-degenerate exciton sublevels in equation (4) under orthorhombic symmetry, which should emit as orthogonal linear dipoles. For simplicity, we assume that the electron and hole Rashba coefficients are equal. The value (0.38 eV Å) required to fit the observed splittings (about 1 meV) is reasonable, lying between those for conventional III–V quantum wells and organic–inorganic perovskites (see Supplementary Information section 1.F). We note that for nanocrystals with tetragonal symmetry, $|\Psi_x\rangle$ and $|\Psi_y\rangle$ in equation (4) remain degenerate (Supplementary Information section 1.E), which explains recently observed two-peak spectra from individual CsPbBr₃ nanocrystals²⁹.

Emitting dipoles that are perpendicular (parallel) to the observation direction should show strong (no) emission. Thus, the intensity from each bright triplet sublevel is explained by both its thermal population and the orientation of the nanocrystal. Single-line spectra (Fig. 3a) arise when the two upper sublevels are unpopulated. Strong linear polarization from this single line (Fig. 3a, inset) supports this interpretation. If the sublevel splitting in this nanocrystal were instead spectrally unresolved, then the line would be unpolarized. From the expected three orthogonal dipoles, we calculated the relative intensity of the photoluminescence peaks and their polarization for arbitrary observation directions (Supplementary Information section 4, Extended Data Fig. 11). We then determined

(Fig. 3d–f) the nanocrystal orientations that are consistent with the spectra and polarizations in Fig. 3a–c. Again, good agreement is obtained.

In Fig. 3g we present the experimental statistics for one-, two- and three-line spectra. One is most common, suggesting that only the lowest sublevel is populated. For the two- and three-line spectra, the measured energy splittings are plotted in Fig. 3h, i. Given three sublevels separated by energies Δ_1 and Δ_2 (Fig. 3i, inset), the average splitting $\bar{\Delta}$ is $0.5(\bar{\Delta}_1 + \bar{\Delta}_2)$, with bars denoting averages. However, two-line spectra can involve any two of the three features, leading to the average $\bar{\Delta}_1/3 + \bar{\Delta}_2/3 + (\bar{\Delta}_1 + \bar{\Delta}_2)/3 = 2(\bar{\Delta}_1 + \bar{\Delta}_2)/3$. We therefore predict a ratio of 1.33 for average measured splittings in two- versus three-line spectra. The experimentally determined ratio of 1.42 ± 0.12 again supports our model.

While we have used cryogenic temperatures to confirm the existence of the bright triplet exciton, its influence on emission remains important at room temperature. Although the splittings are small compared to the thermal energy, the three triplet states (from four sublevels in total) are dipole-allowed and thermally populated, unlike in other nanocrystals^{13–16}. For example, in CdSe nanocrystals only three of eight band-edge sublevels are bright, and these can be poorly populated even at room temperature. This and other effects (Supplementary Information section 5) explain why room-temperature emission from CsPbX₃ perovskite nanocrystals is 20 times faster than in other systems. The emission should be even faster for nanowires and nanoplatelets. Such shapes can further decrease the radiative lifetime, owing to diminished dielectric screening and smaller one- or two-dimensional excitons³⁰.

Although CsPbX₃ nanocrystals are oxidatively stable, their long-term stability may be limited in warm, bright and moist environments without encapsulation to provide thermal and environmental stability. Moreover, the discovery that their lowest exciton is bright reveals criteria for obtaining this phenomenon in other materials. Potential semiconductors should lack inversion symmetry, and one band edge should have *s* symmetry and the other *p*, with the latter affected by strong spin–orbit coupling such that $J_{\text{e,h}} = 1/2$. Finally, the Rashba coefficient for the electron and the hole must be non-zero with the same sign.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 21 June 2017; accepted 9 November 2017.

- Scholes, G. D. & Rumbles, G. Excitons in nanoscale systems. *Nat. Mater.* **5**, 683–696 (2006).
- Hund, F. Concerning the interpretation of complex spectra, especially the elements scandium to nickel. *Z. Phys.* **33**, 345–371 (1925).
- Onodera, Y. & Toyozawa, Y. Excitons in alkali halides. *J. Phys. Soc. Jpn* **22**, 833–844 (1967).
- Nirmal, M. *et al.* Observation of the “dark exciton” in CdSe quantum dots. *Phys. Rev. Lett.* **75**, 3728–3731 (1995).
- Bychkov, Yu. A. & Rashba, E. I. Oscillatory effects and the magnetic susceptibility of carriers in inversion layers. *J. Phys. C* **17**, 6039–6045 (1984).
- Kim, M., Im, J., Freeman, A. J., Ihm, J. & Jin, H. Switchable $S=1/2$ and $J=1/2$ Rashba bands in ferroelectric halide perovskites. *Proc. Natl Acad. Sci. USA* **111**, 6900–6904 (2014).
- Zheng, F., Tan, L. Z., Liu, S. & Rappe, A. M. Rashba spin–orbit coupling enhanced carrier lifetime in CH₃NH₃PbI₃. *Nano Lett.* **15**, 7794–7800 (2015).
- Kepekian, M. *et al.* Rashba and Dresselhaus effects in hybrid organic–inorganic perovskites: from basics to devices. *ACS Nano* **9**, 11557–11567 (2015).
- Mosconi, E., Etienne, T. & De Angelis, F. Rashba band splitting in organohalide lead perovskites: bulk and surface effects. *J. Phys. Chem. Lett.* **8**, 2247–2252 (2017).
- Isarov, M. *et al.* Rashba effect in a single colloidal CsPbBr₃ perovskite nanocrystal detected by magneto-optical measurements. *Nano Lett.* **17**, 5020–5026 (2017).
- Protesescu, L. *et al.* Nanocrystals of cesium lead halide perovskites (CsPbX₃, X=Cl, Br, and I): novel optoelectronic materials showing bright emission with wide color gamut. *Nano Lett.* **15**, 3692–3696 (2015).
- Rainò, G. *et al.* Single cesium lead halide perovskite nanocrystals at low temperature: fast single-photon emission, reduced blinking, and exciton fine structure. *ACS Nano* **10**, 2485–2490 (2016).
- Crooker, S. A., Barrick, T., Hollingsworth, J. A. & Klimov, V. I. Multiple temperature regimes of radiative decay in CdSe nanocrystal quantum dots: Intrinsic limits to the dark-exciton lifetime. *Appl. Phys. Lett.* **82**, 2793–2795 (2003).

- Wuister, S. F., van Houselt, A., de Mello Donegá, C., Vanmaekelbergh, D. & Meijerink, A. Temperature anti-quenching of the luminescence from capped CdSe quantum dots. *Angew. Chem. Int. Ed.* **43**, 3029–3033 (2004).
- Du, H. *et al.* Optical properties of colloidal PbSe nanocrystals. *Nano Lett.* **2**, 1321–1324 (2002).
- Bischof, T. S., Correa, R. E., Rosenberg, D., Dauler, E. A. & Bawendi, M. G. Measurement of emission lifetime dynamics and biexciton emission quantum yield of individual InAs colloidal nanocrystals. *Nano Lett.* **14**, 6787–6791 (2014).
- Shirasaki, Y., Supran, G. J., Bawendi, M. G. & Bulović, V. Emergence of colloidal quantum-dot light-emitting technologies. *Nat. Photon.* **7**, 13–23 (2013).
- Dang, C. *et al.* Red, green and blue lasing enabled by single-exciton gain in colloidal quantum dot films. *Nat. Nanotechnol.* **7**, 335–339 (2012).
- Kim, T. H. *et al.* Full-colour quantum dot displays fabricated by transfer printing. *Nat. Photon.* **5**, 176–182 (2011).
- Bir, G. L. & Pikus, G. E. *Symmetry and Strain-Induced Effects in Semiconductors* Ch. 23 (Wiley, 1974).
- Even, J., Pedesseau, L., Jancu, J.-M. & Katan, C. Importance of spin–orbit coupling in hybrid organic/inorganic perovskites for photovoltaic applications. *J. Phys. Chem. Lett.* **4**, 2999–3005 (2013).
- Kane, E. O. in *Semiconductors and Semimetals* Vol. 1 (eds Willardson, R. K. & Beer, A. C.) 75–100 (Academic Press, 1966).
- Koster, G. F., Dimmock, J. O., Wheeler, R. G. & Statz, H. *Properties of the Thirty-Two Point Groups* (MIT Press, 1963).
- Yaffe, O. *et al.* Local polar fluctuations in lead halide perovskite crystals. *Phys. Rev. Lett.* **118**, 136001 (2017).
- Galland, C. *et al.* Two types of luminescence blinking revealed by spectroelectrochemistry of single quantum dots. *Nature* **479**, 203–207 (2011).
- Cottingham, P. & Brutchey, R. L. On the crystal structure of colloidal prepared CsPbBr₃ quantum dots. *Chem. Commun.* **52**, 5246–5249 (2016).
- Rashba, E. I. & Gugenishvili, G. E. Edge absorption theory in semiconductors. *Sov. Phys. Solid State* **4**, 759–760 (1962).
- Dery, H. & Song, Y. Polarization analysis of excitons in monolayer and bilayer transition-metal dichalcogenides. *Phys. Rev. B* **92**, 125431 (2015).
- Fu, M. *et al.* Neutral and charged exciton fine structure in single lead halide perovskite nanocrystals revealed by magneto-optical spectroscopy. *Nano Lett.* **17**, 2895–2901 (2017).
- Rodina, A. V. & Efros, A. L. Effect of dielectric confinement on optical properties of colloidal nanostructures. *J. Exp. Theor. Phys.* **122**, 554–566 (2016).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank F. Krieg for providing large CsPbBr₃ nanocrystals, S. Yakunin and J. Jagielski for assistance with absolute quantum-yield measurements, and E. Ivchenko, M. Glazov and E. Rashba for discussions. M.A.B., G.R., T.S., M.V.K. and R.F.M. acknowledge the European Union's Horizon-2020 programme through the Marie-Sklodowska Curie ITN network PHONSI (H2020-MSCA-ITN-642656) and the Swiss State Secretariat for Education Research and Innovation (SERI). J.G.M., S.G.L., N.B., J.L.L. and A.L.E. acknowledge support from the US Office of Naval Research (ONR) through the core funding of the Naval Research Laboratory. R.V. was funded by ONR grant N0001416WX01849. A.S. acknowledges support from the Center for Advanced Solar Photonics (CASP), an Energy Frontier Research Center (EFRC) funded by BES, OS, US DOE. D.J.N. and M.V.K. acknowledge partial financial support from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC grant agreement number 339905 (QuaDoPS Advanced Grant) and number 306733 (NANOSOLID Starting Grant), respectively.

Author Contributions This work resulted from ongoing interactions between G.R., M.V.K., D.J.N. and A.L.E. M.A.B., G.R. and T.S. performed the optical experiments, and analysed and interpreted the data with input from D.J.N., R.F.M., P.C.S. and A.L.E. A.L.E. conceived the model and supervised the theoretical research. R.V. calculated the radiative lifetimes and developed the theory for the observed exciton fine structure. R.V. and A.S. developed the four-band model to describe the energy dispersion at the R point and calculated the fine structure using the wavefunction extracted from first-principles calculations. P.C.S. performed the group-theory analysis of the fine structure and selection rules. M.J.M., N.B. and J.L.L. completed the first-principles calculations of the bulk band structures and the band-edge wavefunctions. R.V., P.C.S. and A.L.E. developed the effective exchange and Rashba Hamiltonian that describe the exciton fine structure. J.G.M. and S.G.L. calculated the internal electric fields in spherical and cube-shaped nanocrystals. G.N. prepared the samples and performed electron microscopy under the supervision of M.V.K. A.L.E. and D.J.N. wrote the manuscript with input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to M.V.K. (mvkovalenko@ethz.ch), G.R. (rainog@ethz.ch), D.J.N. (dnorris@ethz.ch) or A.L.E. (sasha.efros@nrl.navy.mil).

Reviewer Information Nature thanks A. Meijerink, E. Rabani and M. Saba for their contribution to the peer review of this work.

METHODS

Chemicals. The following reagents were used to prepare CsPbX₃ nanocrystals: caesium carbonate (Cs₂CO₃, Aldrich, 99.9%), 1-octadecene (ODE, Sigma-Aldrich, 90%), oleic acid (OA, Sigma-Aldrich, 90%), oleylamine (OAm, Acros Organics, 80%–90%), lead chloride (PbCl₂, ABCR, 99.999%), lead bromide (PbBr₂, ABCR, 98%), lead iodide (PbI₂, ABCR, 99.999%), *n*-trioctylphosphine (TOP, Strem, 97%), hexane (Sigma-Aldrich, ≥95%) and toluene (Fisher Scientific, HPLC grade).

Synthesis. The CsPbX₃ (X = Cl, Br or I) and CsPbBr₂Cl nanocrystals were synthesized by fast reaction between Cs-oleate and PbX₂ in the presence of OA and OAm (and TOP in the case of CsPbCl₃ and CsPbBr₂Cl nanocrystals). First, the Cs-oleate was prepared by loading Cs₂CO₃ (0.407 g) into a 50-ml 3-neck flask along with ODE (20 ml) and OA (1.25 ml). The mixture was dried under vacuum for 1 h at 120 °C and then switched to N₂. Because Cs-oleate precipitates out of ODE at room temperature, it must be pre-heated to 100 °C before injection. The ODE, OA and OAm were pre-dried before use by degassing under vacuum at 120 °C for 1 h. For the nanocrystal-forming reaction, 0.376 mmol PbX₂ (X = Cl, Br or I), dried OA (3 ml for PbCl₂, 1 ml for PbBr₂ or 1.5 ml for PbI₂), dried OAm (3 ml for PbCl₂, 1 ml for PbBr₂ or 1.5 ml for PbI₂) and dried ODE (5 ml) were combined in a 25-ml 3-neck flask. For CsPbCl₃, TOP (1 ml) was also added. The mixture was then degassed for 10 min under vacuum at 120 °C, and the flask was filled with N₂ and heated to 200 °C. Cs-oleate (0.8 ml from the stock solution prepared as described above) was injected swiftly when 200 °C was reached. After 10 s the reaction was stopped by cooling the reaction system with a water bath. The solution was centrifuged (4 min, 13,750g) and the supernatant discarded. Hexane (0.3 ml) was added to the precipitate to disperse the nanocrystals and the mixture was then centrifuged again. The precipitate obtained was re-dispersed in 3 ml toluene and centrifuged (2 min, 2,200g). The supernatant was separated from the precipitate, filtered and used for our investigations. For CsPbBr₂Cl, 0.094 mmol PbCl₂, 0.282 mmol PbBr₂, dried OA (1.5 ml), dried OAm (1.5 ml), TOP (1 ml) and dried ODE (5 ml) were loaded into a 25-ml 3-neck flask and the same protocol was followed.

Sample preparation. For single-nanocrystal spectroscopy, the colloidal dispersions from the above syntheses were diluted to nanomolar concentrations in solutions of 3-wt% polystyrene in toluene. This dispersion was then spin-casted at 5,000 r.p.m. onto intrinsic crystalline Si wafers with a 3-μm-thick thermal-oxide layer. For ensemble measurements, the undiluted nanocrystal dispersions were drop-casted on glass substrates. For photoluminescence quantum-yield measurements, 0.1 ml of the colloidal dispersion was mixed with 0.1 ml of a 5-wt% solution of poly(methyl methacrylate) (PMMA, molecular weight of 495,000) in toluene.

Optical characterization. All optical measurements of single nanocrystals were performed in a self-built micro-photoluminescence (μ-PL) set-up. The samples were mounted on xyz nano-positioning stages inside an evacuated liquid-helium flow cryostat and cooled down to a targeted temperature of 5 K (see Fig. 3 caption). Single nanocrystals were excited by means of a fibre-coupled excitation laser at an energy of 3.06 eV with a repetition rate of 40 MHz and a pulse duration of 50 ps. The excitation beam was sent through a linear polarizer and a short-wavelength-pass filter before being directed towards the sample by a dichroic beam splitter. Typical power densities used to excite single nanocrystals were 2–120 W cm⁻². Assuming an absorption cross section³¹ of 8×10^{-14} cm², these power densities yield 0.0057–0.34 excitons per nanocrystal per pulse³². For both excitation and detection, a long-working-distance 100× microscope objective with a numerical aperture of 0.7 was used. The nearly Gaussian excitation spot had a $1/e^2$ diameter of 1.4 μm. The emission was filtered using a long-pass filter and dispersed by a 0.75-m monochromator with a grating of 1,800 lines mm⁻¹ before detection with a back-illuminated, cooled charge-coupled device camera. For polarization-dependent measurements, a liquid-crystal retarder was used to compensate for retardation effects in the set-up. For photoluminescence lifetime and time-tagged time-resolved (TTTR3) single-photon-counting measurements, we filtered the emission with a suitable tunable bandpass filter either to measure only the excitonic photoluminescence decay or to correlate excitonic and trionic emission

intensities and decay times with a time-correlated single-photon-counting system with nominal time resolution of 30 ps.

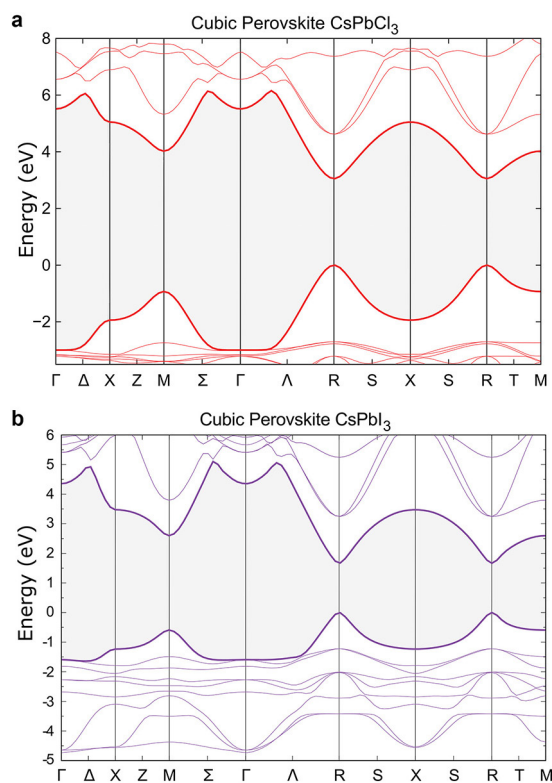
Ensemble measurements were performed in an exchange-gas cryostat at 5 K. Here, the samples were excited with a frequency-doubled Ti:sapphire femtosecond pulsed laser with a repetition rate of 80 MHz at 3.1 eV. Optical power densities were below 3 W cm⁻². The emitted light was dispersed by a grating of 150 lines per mm within a 300-mm-focal-length spectrograph and detected by a streak camera with 2-ps resolution. Absolute photoluminescence quantum-yield measurements at room temperature were performed on a Quantaurus QY (C11347-11, Hamamatsu).

Band-structure calculations. Figure 1b and Extended Data Fig. 1 show calculated band structures for CsPbBr₃, CsPbCl₃ and CsPbI₃. We assume that these materials exist in the cubic perovskite structure with a lattice constant of 5.865 Å, 5.610 Å and 6.238 Å, respectively³³. The electronic structure of these crystals was determined using the Vienna *Ab-initio* Simulation Package (VASP)^{34–36} with projector-augmented wavefunctions³⁷. Our initial calculations used the PBEsol³⁸ generalized gradient approximation, and included spin–orbit coupling. We used an energy cut-off of 400 eV and a Γ -centred *k*-point grid of $6 \times 6 \times 6$, which yield 40 *k*-points in the irreducible Brillouin zone.

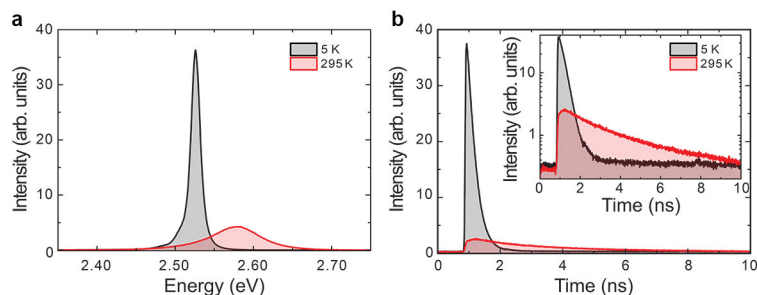
As expected, standard density functional theory (DFT) underestimates the bandgap in these materials substantially. Accordingly, we used a modified version of the Heyd–Scuseria–Ernzerhof ‘HSE06’ hybrid functional³⁹, which mixes exact Hartree–Fock exchange with conventional DFT. We initially started with 25% mixing, and planned to adjust the mixing to match the observed bandgap. However, this was not possible, even with 45% Hartree–Fock in the calculation for CsPbBr₃. This initial mixing produced a bandgap of 1.4 eV, far smaller than the experimentally determined gap of 2.8 eV. Rather than using even higher mixing, or even a full-scale Hartree–Fock calculation, we instead added a scissors operator to adjust the bandgap to the experimental result. We found that the electron and hole masses were nearly unchanged with Hartree–Fock mixing, leading us to believe that this technique still provides the correct physics. Further confirmation was provided by conducting *G₀W₀* calculations (also with VASP) on top of the PBE results. For this approach, we used a plane-wave energy cut-off of 600 eV, a 150-eV energy cut-off for the response functions, 1,894 unoccupied states, spin–orbit coupling, and ‘GW’ pseudopotentials including all semi-core electrons. Although these calculations yielded bandgaps that were in closer agreement with the experiments (1.96 eV for CsPbI₃, 2.36 eV for CsPbBr₃ and 3.27 eV for CsPbCl₃), other aspects of the band structure remained virtually unchanged.

Data availability. All data generated or analysed during this study are included in the published article (and its Supplementary Information).

- Yakunin, S. *et al.* Low-threshold amplified spontaneous emission and lasing from colloidal nanocrystals of caesium lead halide perovskites. *Nat. Commun.* **6**, 8056 (2015).
- Smyder, J. A. *et al.* The influence of continuous vs. pulsed laser excitation on single quantum dot photophysics. *Phys. Chem. Chem. Phys.* **16**, 25723–25728 (2014).
- Mehl, M. J. *et al.* The AFLOW library of crystallographic prototypes: part 1. *Comput. Mater. Sci.* **136** (Suppl.), S1–S828 (2017).
- Kresse, G. *Ab Initio Molekular Dynamik für Flüssige Metalle*. PhD thesis, Technische Universität Wien (1993).
- Kresse, G. & Hafner, J. *Ab initio* molecular dynamics for open-shell transition metals. *Phys. Rev. B* **48**, 13115–13118 (1993).
- Kresse, G. & Hafner, J. *Ab initio* molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251–14269 (1994).
- Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
- Perdew, J. P. *et al.* Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008); erratum 102, 039902 (2009).
- Krukau, A. V., Vydrov, O. A., Izmaylov, A. F. & Scuseria, G. E. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *J. Chem. Phys.* **125**, 224106 (2006).



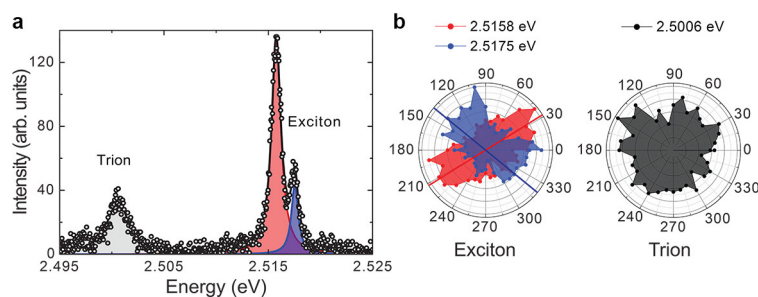
Extended Data Figure 1 | Electronic structure for CsPbCl_3 and CsPbI_3 perovskites. **a**, Calculated band structure of cubic perovskite CsPbCl_3 . **b**, Calculated band structure of cubic perovskite CsPbI_3 . See Methods for details about the calculations.



Extended Data Figure 2 | Measurements to estimate the low-temperature quantum yield for our CsPbBr₂Cl nanocrystals.

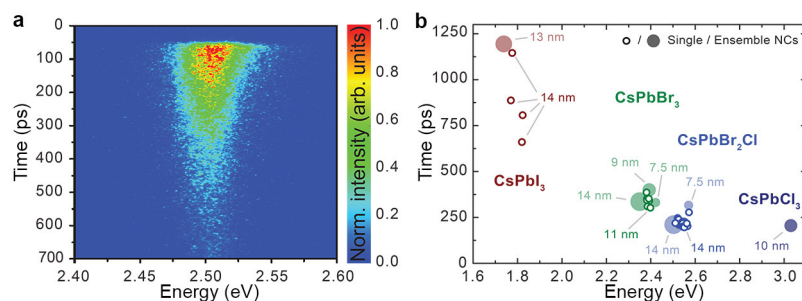
a, b, Photoluminescence spectra (**a**) and decays (**b**) for CsPbBr₂Cl nanocrystals ($L = 14 \pm 1$ nm) embedded in a PMMA film at 295 K (red) and 5 K (black). Data for the two temperatures are plotted on the same intensity scale. For the same sample, a calibrated integrating sphere was used to measure the photoluminescence quantum yield at 295 K ($43\% \pm 1\%$). To obtain the quantum yield at 5 K, the photoluminescence and optical absorption for several spots at 295 K and 5 K under constant

weak excitation (at 3.06 eV) were measured. The photoluminescence increased substantially, as seen in both the spectra and decay signal, whereas the absorption stayed nearly constant (data not shown). From these results, the quantum yield at 5 K was estimated to be $88\% \pm 14\%$. The photoluminescence decays in **b** are plotted on both a linear and a logarithmic (inset) intensity scale, with decay times of 1.60 ns (295 K) and 0.23 ns (5 K). The decrease in decay time at low temperature is clearly accompanied by an increase in the total emitted intensity (area under the decay traces).



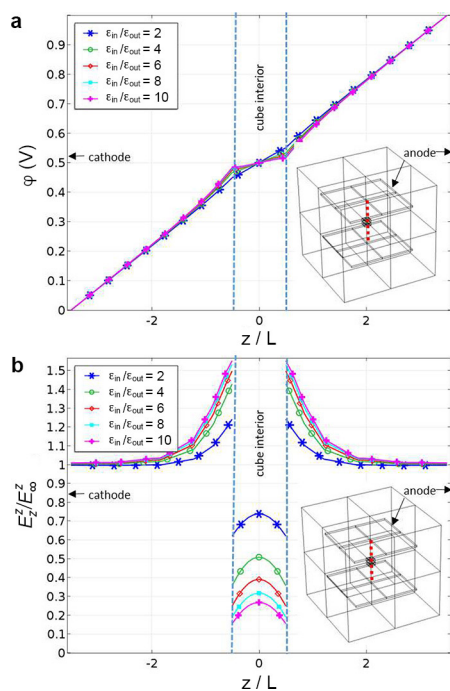
Extended Data Figure 3 | Exciton and trion emission from an individual CsPbBr₂Cl nanocrystal. **a**, Photoluminescence spectrum of a single CsPbBr₂Cl nanocrystal, showing exciton peaks at 2.5158 eV (red) and 2.5175 eV (blue) and a trion peak (black) that is redshifted by 15–17 meV. The targeted temperature was 5 K (see Fig. 3 caption). **b**, Polarization properties of the exciton (left) and trion (right) emission peaks. The

normalized area of a Lorentzian-peak fit for the two exciton peaks (red and blue) and the trion peak (black) are shown as a function of the linear polarizer angle (placed in front of the spectrograph). Both exciton peaks show a dominantly linear polarization, with the main axis indicated by the blue and red lines. The trion emission is unpolarized. See Supplementary Section 4 for further discussion.

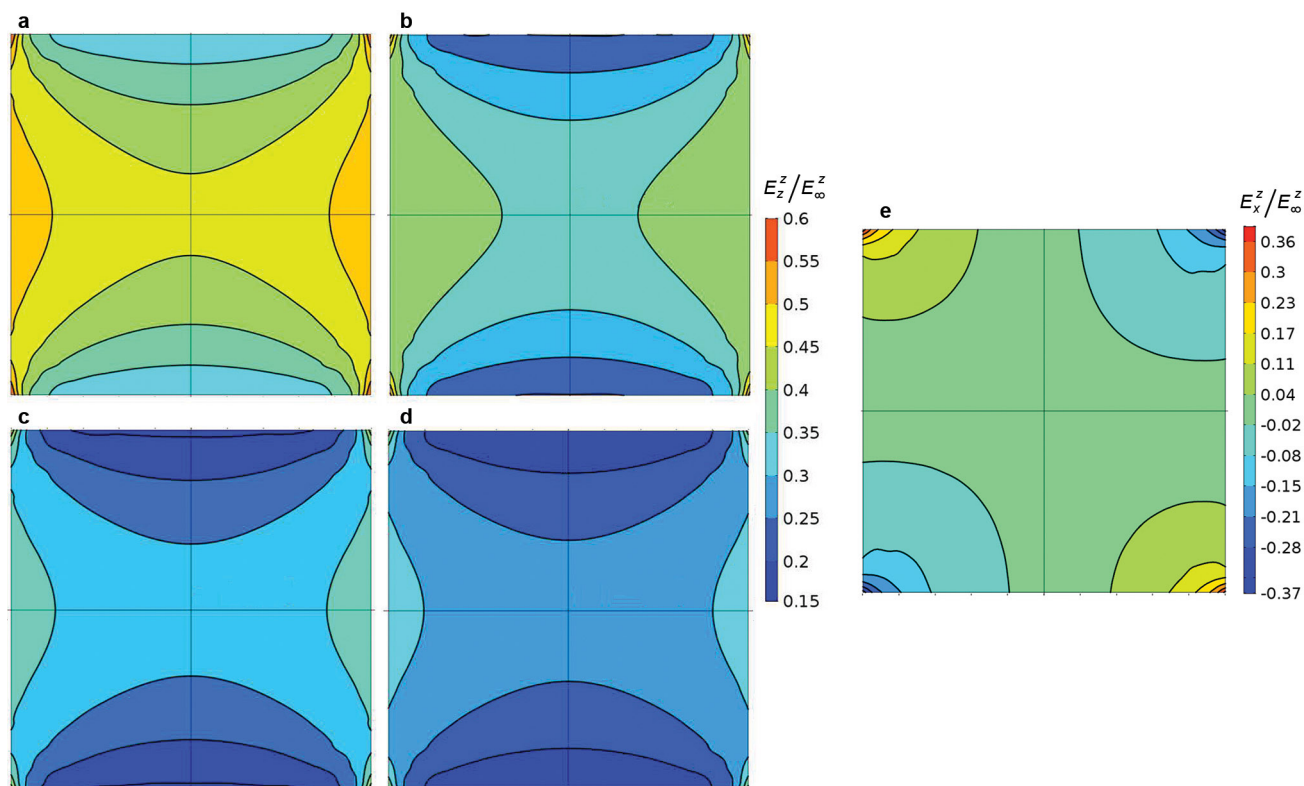


Extended Data Figure 4 | Composition-dependent ensemble photoluminescence decay measurements of lead halide perovskite nanocrystals. **a**, Typical streak-camera measurement of the photoluminescence from an ensemble of CsPbBr₂Cl nanocrystals at 5 K. In this example, the nanocrystals have $L = 14$ nm. The emission peak is centred at 2.51 eV, and the exponential decay time is 210 ps, as extracted by summing over all energies, which is in good agreement with the results for single CsPbBr₂Cl nanocrystals of the same size. The ensemble decay spectrum is slightly asymmetric (being faster at higher energies), which might originate from the activation of an energy-transfer process from

smaller to larger nanocrystals. To account for this effect, we considered only the long component of the decay curve. **b**, Photoluminescence lifetimes at 5 K extracted for ensemble samples of nanocrystals (NCs) of various compositions and sizes (as labelled). The ensemble data (solid circles) are compared with single-nanocrystal measurements (open circles). The good agreement between the two datasets is further evidence that the measured single-nanocrystal photoluminescence decays are due to fast exciton radiative lifetimes and not to trions, because the ensemble data are acquired at very low excitation power, at which photo-generated charging is not observed.

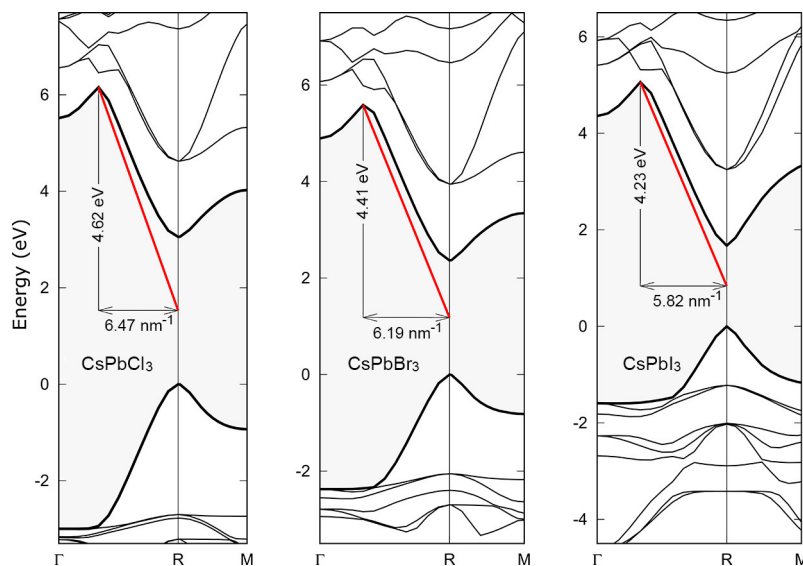


Extended Data Figure 5 | Calculation of the interior electric field in cube-shaped nanocrystals. a, Line plot of the electric potential ϕ along the centre line between the capacitor plates (see inset and Supplementary Information section 3.B). **b,** Line plot of the normalized electric-field magnitude E_z^z/E_∞^z along the centre line between the capacitor plates (see inset).



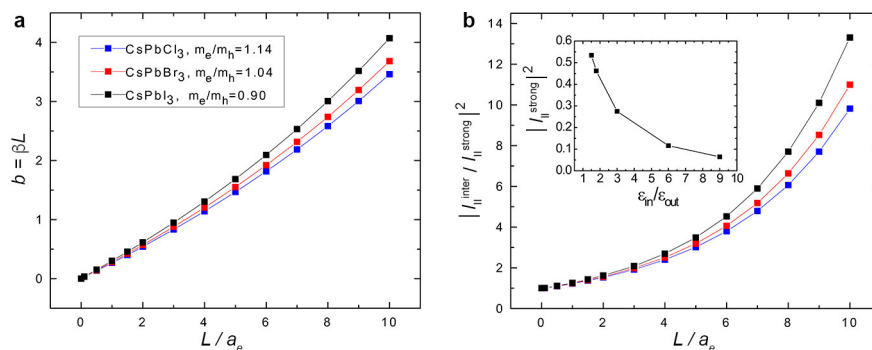
Extended Data Figure 6 | Contour plots of normalized electric-field magnitude across a cube-shaped nanocrystal. a–d, Contour plots of E_z^z/E_∞^z for four different ratios (4, 6, 8 and 10, respectively) of the dielectric constant inside the nanocrystal (ε_{in}) and of the surrounding medium (ε_{out}) (see Supplementary Information section 3.B). The plots depict the x – z mid-plane of the cube and are valid for the symmetry-

equivalent y – z mid-plane. **e,** Contour plot of E_x^z/E_∞^z on the x – z mid-plane of the cube for $\varepsilon_{\text{in}}/\varepsilon_{\text{out}} = 9$. The E_y^z/E_∞^z distribution on the y – z mid-plane is identical. In all panels, the z direction is vertical and the perturbations near the corners of the plots are artefacts of the interpolation resolution used by the software that we used to construct them.



Extended Data Figure 7 | Extraction of the Kane energy E_p for the lead halide perovskites. The Kane energy, defined according to equation (5), can be extracted for CsPbCl_3 , CsPbBr_3 and CsPbI_3 from the band

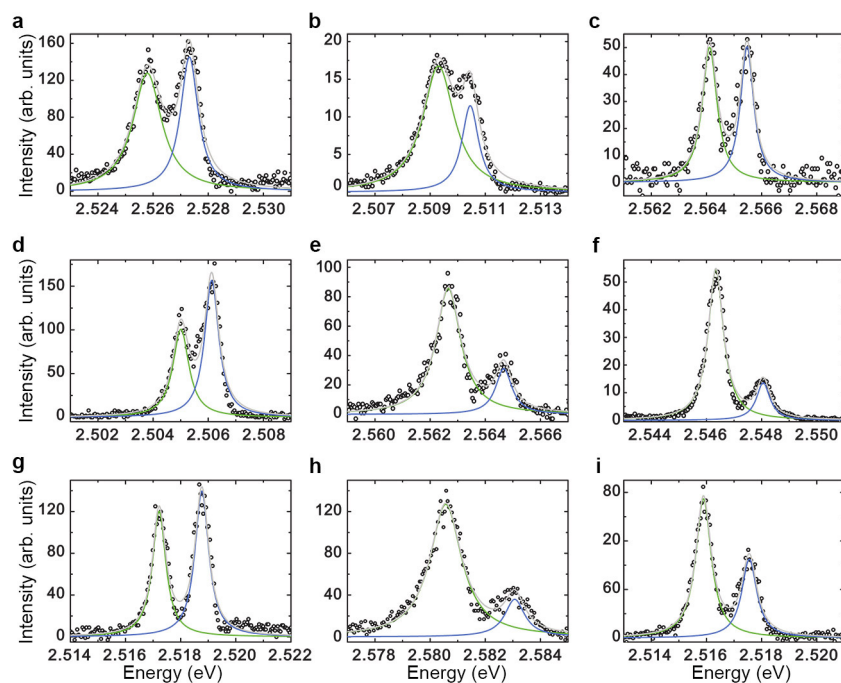
structures presented in Fig. 1b and Extended Data Fig. 1 near the band edges. The slope of the red line is used, according to a procedure described in Supplementary Information section 1.B.



Extended Data Figure 8 | Variational calculations related to the determination of the exciton radiative lifetime in cube-shaped nanocrystals within the intermediate-confinement regime.

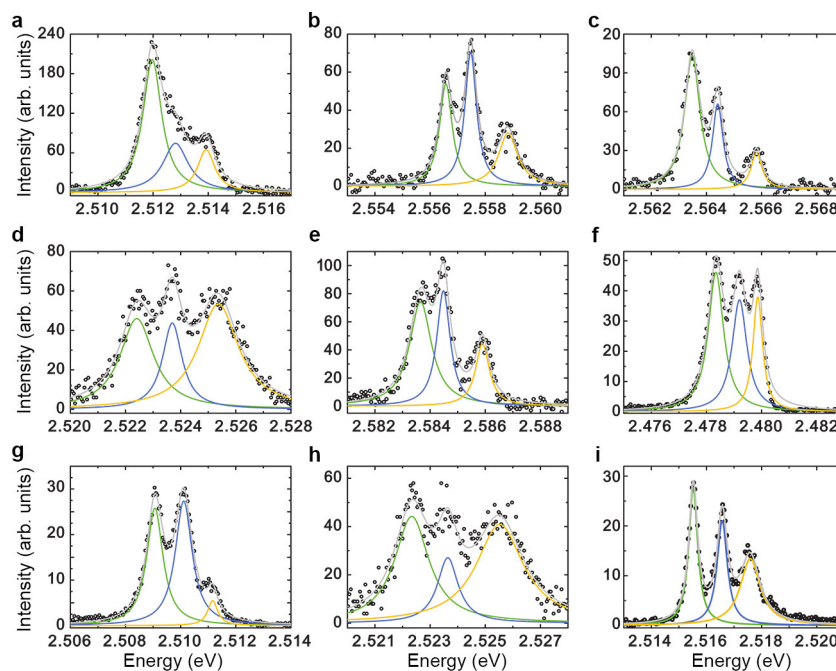
a, b, Dimensionless electron-hole correlation constant ($b = \beta L$, where β is the value of the variational parameter that minimizes the energy; **a**) and the square modulus of the ratio of $I_{||}$ for intermediate and strong

confinement (**b**) as a function of the size of the nanocrystal relative to the Bohr radius of an electron (L/a_e), for the three materials studied; m_e and m_h are the electron and hole effective masses, respectively. The inset in **b** shows the square modulus of $I_{||}$ in the strong-confinement regime for several different dielectric constants, $\epsilon_{\text{in}}/\epsilon_{\text{out}}$. See Supplementary Information section 3.D for details.

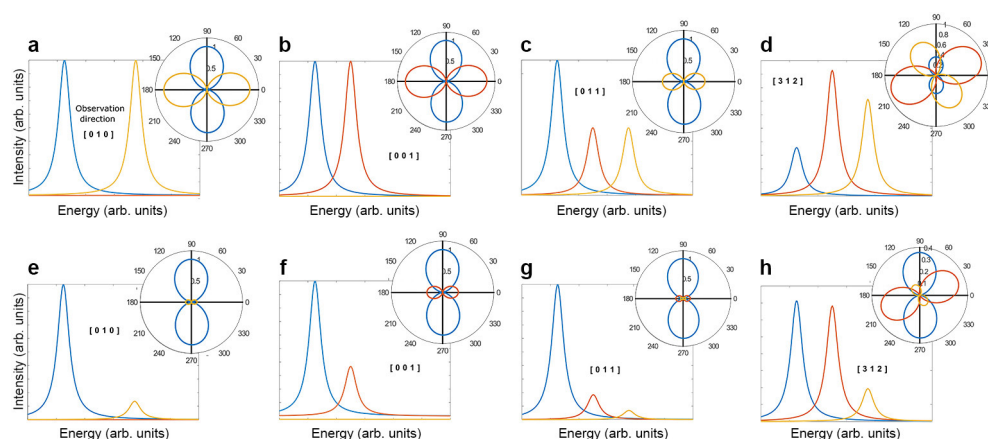


Extended Data Figure 9 | Representative two-peak spectra for individual CsPbBr_2Cl nanocrystals. a–i, Photoluminescence spectra of different single nanocrystals that exhibit two emission peaks at a targeted temperature of 5 K (see Fig. 3 caption). Each spectrum was recorded with a linear polarizer in the detection path. Therefore, the relative intensities

displayed cannot be used to determine the relative (potentially thermal) population within the fine-structure multiplet. The linear polarizer was used here because it can be rotated to resolve all spectral features. Without the polarizer, the low-energy peak typically dominates in intensity.



Extended Data Figure 10 | Representative three-peak spectra for individual CsPbBr₂Cl nanocrystals. Details as for Extended Data Fig. 9, but for nanocrystals that exhibit three emission peaks.



Extended Data Figure 11 | Predicted exciton spectra and polarization properties for individual perovskite nanocrystals. The plots show the expected exciton fine structure in photoluminescence spectra from three orthogonal dipoles of the lowest-energy exciton. The dipoles are oriented along the orthorhombic symmetry axes. The insets show the emission probability for the dipoles as a function of the polarization angle. **a–d**, Expected fine structure for observation in the [010], [001], [011] and [312] directions with respect to the orthorhombic symmetry axes,

respectively. The temperature effect on the population of the sublevels is not considered (that is, the populations of the sublevels are assumed to be equal). **e–h**, As in **a–d**, but taking the temperature effect on the population of the sublevels into consideration. The temperature is assumed to be comparable to the fine-structure splitting: $k_B T \approx \Delta_1 = \Delta_2$, where k_B is the Boltzmann constant and T is temperature. See Supplementary Information section 4 for further details.

Fire frequency drives decadal changes in soil carbon and nitrogen and ecosystem productivity

Adam F. A. Pellegrini¹, Anders Ahlström^{1,2}, Sarah E. Hobbie³, Peter B. Reich^{4,5}, Lars P. Nieradzik⁶, A. Carla Staver⁷, Bryant C. Scharenbroch⁸, Ari Jumpponen⁹, William R. L. Anderegg¹⁰, James T. Randerson¹¹ & Robert B. Jackson^{1,12}

Fire frequency is changing globally and is projected to affect the global carbon cycle and climate^{1–3}. However, uncertainty about how ecosystems respond to decadal changes in fire frequency makes it difficult to predict the effects of altered fire regimes on the carbon cycle; for instance, we do not fully understand the long-term effects of fire on soil carbon and nutrient storage, or whether fire-driven nutrient losses limit plant productivity^{4,5}. Here we analyse data from 48 sites in savanna grasslands, broadleaf forests and needleleaf forests spanning up to 65 years, during which time the frequency of fires was altered at each site. We find that frequently burned plots experienced a decline in surface soil carbon and nitrogen that was non-saturating through time, having 36 per cent (± 13 per cent) less carbon and 38 per cent (± 16 per cent) less nitrogen after 64 years than plots that were protected from fire. Fire-driven carbon and nitrogen losses were substantial in savanna grasslands and broadleaf forests, but not in temperate and boreal needleleaf forests. We also observe comparable soil carbon and nitrogen losses in an independent field dataset and in dynamic model simulations of global vegetation. The model study predicts that the long-term losses of soil nitrogen that result from more frequent burning may in turn decrease the carbon that is sequestered by net primary productivity by about 20 per cent of the total carbon that is emitted from burning biomass over the same period. Furthermore, we estimate that the effects of changes in fire frequency on ecosystem carbon storage may be 30 per cent too low if they do not include multidecadal changes in soil carbon, especially in drier savanna grasslands. Future changes in fire frequency may shift ecosystem carbon storage by changing soil carbon pools and nitrogen limitations on plant growth, altering the carbon sink capacity of frequently burning savanna grasslands and broadleaf forests.

Fire regimes have been altered by changes in climate and land use, and are predicted to change further as temperatures rise and populations grow^{1–3}. In consequence, the response of ecosystems to long-term alterations in fire frequency—that is, either more frequent burning or fire suppression—will be essential to the future of the terrestrial carbon sink^{3,6}. Although carbon fluxes to the atmosphere from combusting plant biomass have been well characterized⁷, uncertainties remain concerning the responses of soil carbon and nutrient pools^{4,5}, which also regulate plant primary productivity⁸.

On the one hand, increased burning may decrease soil organic matter, as repeated burning reduces organic inputs to soils and leads to declines in soil carbon (C) and nutrients^{9–11}. On the other hand, increased burning may enrich C and nutrient concentrations in soils by promoting the establishment of more-productive plant species¹² and the leaching of ash downwards into soils¹³. Observations generally

illustrate that single fires deplete pools of C and nutrients in the surface litter layer and, in some cases, in shallow organic horizons^{14,15}. Critically, however, studies that document changes in soils over short timescales or in response to a single fire (see, for example, refs 13, 14) offer limited insight into long-term changes in the larger mineral soil pools as a result of shifting fire regimes, particularly in soils below the top few centimetres; such soils are generally not subject to direct consumption¹⁶ and are influenced more by fire-induced changes in plant inputs and microbial activity^{10,17}. Thus generalized long-term effects of changes in fire frequencies on soil C and nitrogen, and on their controlling mechanisms, remain unclear, with contrasting results observed in studies of different regions or ecosystems^{10,11,17}.

A lack of consensus on the long-term response of soils to fire limits our ability to predict how vegetation productivity may change as fire alters soil nutrient availability. Over the short-term, single fires can stimulate plant productivity¹⁸; however, over the longer-term, potential declines in soil nutrients with increased fire frequency^{9–11} have been hypothesized to suppress productivity, although long-term evidence for this effect is limited¹¹. These interactions may determine whether fire reduces ecosystem C storage by depleting soil C and nutrients, which may reduce plant growth and turnover, further constraining C storage in the ecosystem (Supplementary Fig. 1).

Here, we evaluate these interactions by examining how long-term differences in fire frequency alter soil C and nutrients and accompanying shifts in plant productivity, using three approaches. First, we use a meta-analysis of data from 48 sites worldwide (Fig. 1a) to test how frequent burning alters soil C and nutrients over time spans as long as 65 years. We then evaluate our results using an independent dataset from 16 additional field sites, which were not replicated at the site scale (and thus were not included in the meta-analysis), but collectively are valuable given the high number of sites and standardized data collection. Finally, we use our results to validate an individual-based dynamic global vegetation model (the DGVM LPJ-GUESS-BLAZE) for quantifying the effect of fire-driven nutrient losses on vegetation productivity and the degree to which soils contribute to ecosystem-level changes in C.

The sites included in the meta-analysis compared the effects of changes in long-term fire frequencies on C and nutrients in the upper soil layer (0–20 cm depth); the average treatment length was 30 years and ranged from 9–65 years. Sites generally contained plots that either experienced elevated fire frequency (4.3 ± 0.6 times more than the estimated historical mean for that ecosystem, calculated over the length of the study) or were protected from fire (complete fire exclusion in all but one case), which we refer to hereafter as ‘elevated’ and ‘protected’ treatments, respectively (see Supplementary Information). Sites covered

¹Department of Earth System Science, Stanford University, Stanford, California 94305, USA. ²Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden.

³Department of Ecology, Evolution, and Behavior, University of Minnesota, St Paul, Minnesota 55108, USA. ⁴Department of Forest Resources, University of Minnesota, St Paul, Minnesota 55108, USA. ⁵Hawkesbury Institute for the Environment, Western Sydney University, Sydney, New South Wales, Australia. ⁶Centre for Environmental and Climate Research, CEC, Lund University, Lund, Sweden.

⁷Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06520, USA. ⁸College of Natural Resources, University of Wisconsin–Stevens Point, Stevens Point, Wisconsin 54481, USA. ⁹Division of Biology, Kansas State University, Manhattan, Kansas 66506, USA. ¹⁰Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA.

¹¹Department of Earth System Science, University of California–Irvine, Irvine, California 92697, USA. ¹²Woods Institute for the Environment and Precourt Institute for Energy, Stanford University, Stanford, California 94305, USA.

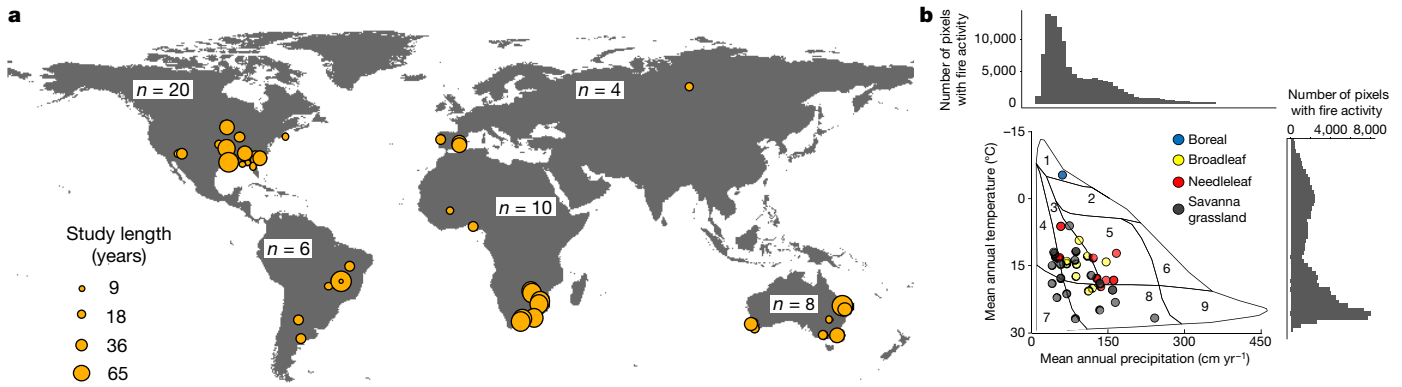


Figure 1 | Distribution of study sites. **a**, Geographical distribution of sites ($n = 48$), with dot size representing study duration. **b**, Climatic distribution of sites. Bottom left, vegetation types are indicated by different colours plotted over a modified diagram of Whittaker's biomes³⁰ (1, tundra; 2, boreal forest; 3, woodland/shrubland; 4, temperate grassland/desert; 5, temperate forest; 6, temperate rainforest; 7, subtropical desert;

8, tropical forest and savanna; 9, tropical rainforest). Dots are slightly transparent to allow overlap to be visualized. The histograms above and to the right illustrate the frequency distribution of global fire activity for a given climatic condition. Fire activity was determined using gridded maps of mean fire occurrence taken from the global fire emissions database 4 with small fires (GFED4s)⁷.

a broad range of mean annual temperature (-5 – 27°C) and precipitation (410 – $2,410\text{ mm yr}^{-1}$) (Fig. 1b and Supplementary Fig. 2). To evaluate whether fire effects depended on plant communities, we categorized sites on the basis of the dominant plant functional type into savanna grasslands, broadleaf forests and needleleaf forests. Statistical significance was evaluated using mixed-effects models of the logarithmic response ratio (natural logarithm of the quotient between elemental concentration in elevated and protected plots), weighted by site replication and variance¹⁹ (Supplementary Fig. 3).

We found that elevated fire frequencies substantially decreased total soil C and nitrogen (N) concentrations globally, with the largest effects observed in broadleaf forests and savanna grasslands. Averaged across all sites, vegetation types, and treatment lengths, higher fire frequencies reduced the concentrations of total soil C and N concentrations by

12.1% (confidence interval $\pm 10.2\%$; $P = 0.02$) and 10.4% ($\pm 10.0\%$; $P = 0.04$), respectively, compared with plots protected from fire (Fig. 2a, b and Supplementary Table 1; 30-year mean treatment length). Within vegetation types, fires had strong depletion effects on soils in both broadleaf forests (27% less C and 25% less N in elevated versus protected plots; $P < 0.001$ and $P = 0.02$, respectively) and savanna grasslands (21% less C and N in elevated versus protected plots; $P < 0.001$ for each; Fig. 2a, b and Supplementary Table 1). By contrast, soil C and N in needleleaf forests increased by 26% and 21%, respectively, in elevated compared with protected plots ($P < 0.001$ for each; Fig. 2a, b and Supplementary Table 1). The different responses that we observed in needleleaf forests were unlikely to be caused by climatic variables or study design, given that, in our dataset, there were no differences between sites in temperate needleleaf forests and those in savanna

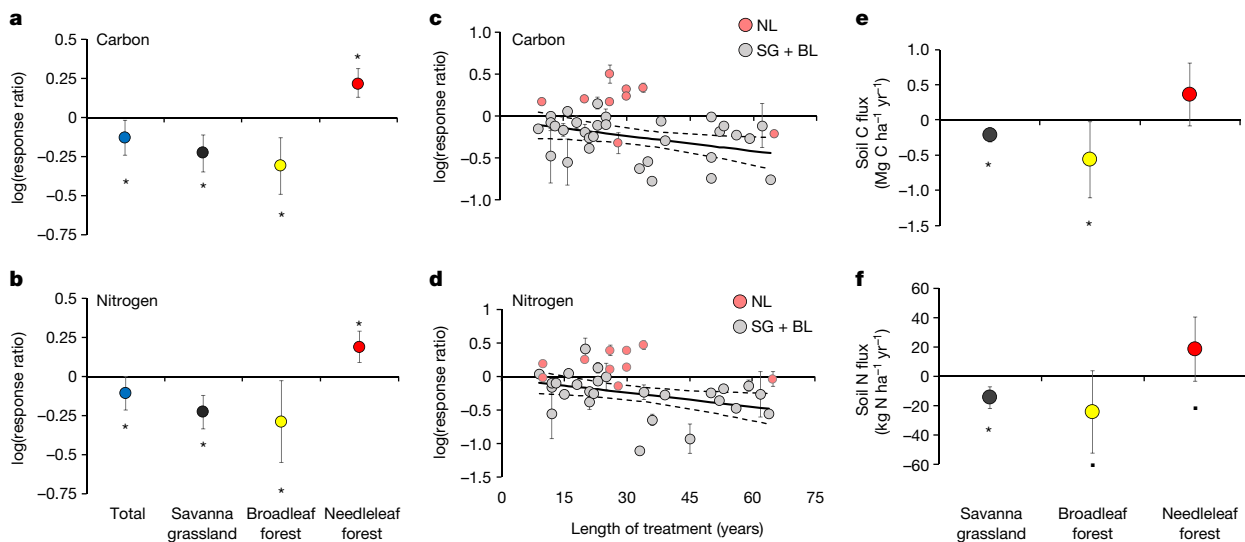


Figure 2 | Effects of fire on soil carbon and nitrogen across ecosystems and over time. **a**, **b**, Logarithmic response ratios of the concentrations of C (**a**, $n = 41$) and N (**b**, $n = 38$) for the total dataset compiled and partitioned into different vegetation types (see Supplementary Tables 1 and 2 for statistics). The response ratio is defined as the concentration of C or N in elevated plots divided by the concentration in protected plots. **c**, **d**, Regressions between the response ratios of C (**c**, $n = 31$) or N (**d**, $n = 27$) and the length of time during which plots experienced contrasting fire frequencies, fitted for savanna grasslands (SG) plus

broadleaf forests (BL) using a meta-regression. Pink dots represent data from needleleaf forests (NL), which were not used in the regression. **e**, **f**, Total fluxes of C (**e**) and N (**f**), determined as the absolute rate of change in soil C or N between the fire frequency treatments (negative values indicate losses under frequent burning). Dashed lines represent 95% confidence intervals (for **c**, **d**) and error bars represent either 95% confidence intervals (for **a**, **b**, **e**, **f**) or the variance around the logarithmic response ratio (for **c**, **d**; see ref. 19), with asterisks indicating significance at $P < 0.05$ and dots at $P < 0.10$ (Supplementary Tables 4 and 5).

grasslands and broadleaf forests in climatic conditions, sampling depths, or fire frequency in elevated plots (Supplementary Tables 2 and 3). The effect of fire in boreal needleleaf forests, which differ substantially in climate compared with the other vegetation types (Fig. 1b), was similar to its effect in temperate needleleaf forests (see Supplementary Information). N stocks in mineral soils tended to increase with more frequent burning ($r^2 = 0.24$, $P = 0.058$), whereas C stocks displayed no trend (Supplementary Fig. 4).

In savanna grasslands and broadleaf forests, the severity of fire-driven losses of soil C and N increased significantly with the length of time for which plots experienced altered fire frequencies. Soils in elevated plots were estimated to have 36% and 38% less C ($P = 0.026$) and N ($P = 0.022$), respectively, than those in protected plots after 64 years (the maximum duration in savanna grassland and broadleaf forest sites; Fig. 2c, d and Supplementary Table 4). Furthermore, for both C and N, the difference between elevated and protected plots differed significantly ($P < 0.05$) only after 18 years of contrasting fire frequencies, highlighting that effects emerge over decadal timescales. By contrast, the responses in needleleaf forests were unchanged with increasing duration of fire treatment ($P > 0.5$ for C and N; Fig. 2c, d).

To further evaluate the generality of our global meta-analysis, we analysed an independent dataset from a network of 16 additional field experiments across the southeastern United States (see Supplementary Information). Of those sites that experienced different fire frequencies for a duration sufficient to detect a potential effect, 83% showed declines in C and 67% showed declines in N with frequent burning; elevated sites had on average 13% and 11% lower C and N, respectively, than did protected plots (Supplementary Fig. 5). Considering the shorter average length of time that these plots experienced different fire frequencies (22 years), the mean responses are consistent with results from the global meta-analysis regression between C and N losses and study length ($17\% \pm 10\%$ for C and N; Supplementary Fig. 5).

To determine changes in total stocks of C and N in response to fire alterations, we combined elemental concentrations with soil bulk densities to a standardized depth of 10 cm, and normalized stock changes to an annual rate from the meta-analysis. The subset of studies that did not provide bulk density data required values to be extrapolated on the basis of soil texture or by using the mean value (see Supplementary Information). Plots exposed to elevated fire frequencies experienced large average losses of soil C and N stocks relative to protected plots in savanna grasslands ($-0.21 \text{ Mg C ha}^{-1} \text{ yr}^{-1}$ and $-14.5 \text{ kg N ha}^{-1} \text{ yr}^{-1}$; $P < 0.001$ for both) and broadleaf forests ($-0.57 \text{ Mg C ha}^{-1} \text{ yr}^{-1}$ and $-24.3 \text{ kg N ha}^{-1} \text{ yr}^{-1}$; $P < 0.05$ and $P < 0.1$, respectively) (Fig. 2e, f and Supplementary Table 5). By contrast, there was no change in soil C stocks, and a marginally significant enrichment of soil N, in needleleaf forests in elevated plots ($+18.4 \text{ kg N ha}^{-1} \text{ yr}^{-1}$; $P < 0.1$) (Fig. 2e, f and Supplementary Table 5).

We found little evidence that increased fire frequencies depleted other elements besides C and N. Averaged across all sites, surface mineral soils in elevated plots showed no change in concentrations of phosphorus (P) relative to protected plots (Fig. 3a and Supplementary Table 6), but they were enriched in calcium ($+52\%$; $P < 0.0001$) and potassium ($+13\%$; $P = 0.02$) (Supplementary Table 6). The duration of fire frequency alterations influenced the direction and significance of results only for soil P. Concentrations of P were initially enriched in the elevated plots after a decade of burning ($+51\%$; $P = 0.01$), but this effect disappeared after about 30 years of frequency alterations (Fig. 3b and Supplementary Table 7). Longer-term studies are needed to determine whether exposure to fire will deplete soil P because of enhanced erosion; however, of the five sites in our analysis that experienced more than 50 years of altered fire frequencies, only one was depleted in P. The lack of P, potassium and calcium losses following long-term changes in fire frequency is consistent with the hypothesis that their higher oxidation temperatures and/or soil sorption capacities decrease losses during frequent burning compared with C and N²⁰.

Changes in fire frequency can also alter plant-available nutrients. Across the global dataset, elevated-frequency plots had 25% lower concentrations of inorganic N (the main form of N available to plants) relative to protected plots ($P < 0.0001$), with a positive correlation found between total N and inorganic N response ratios (Supplementary Fig. 6). By contrast, there was no significant effect of fire frequency on concentrations of inorganic P (the main form of P available to plants). The responses of inorganic P and total P were positively correlated (Supplementary Fig. 7). Our data clearly show that the observed significant increases in inorganic N immediately following fires (see, for example, ref. 14) are transient, and often reverse with repeated burning.

Given the importance of soil N for sustained productivity, we next evaluated the degree to which N losses might constrain plant net primary productivity (NPP), potentially restricting C uptake. To do so, we simulated the effect of fire on ecosystem C and N by using the DGVM LPJ-GUESS²¹ with the process-based fire module BLAZE (see Supplementary Information). For each study site, we simulated ecosystem dynamics for the period 1950–2013, using fire frequencies, climate, and N deposition specific to each site, as well as changes in global CO_2 concentrations (see Supplementary Information).

Like our empirical data, the model showed losses (albeit smaller ones) of total soil C and N in response to frequent burning in both broadleaf forests and savanna grasslands (Supplementary Figs 8 and 9). However, the model also simulated net losses of soil C and N from needleleaf sites, unlike the empirical data (Supplementary Fig. 10), illustrating the need for further model development and additional data. In broadleaf forests and savanna grasslands, simulated declines in total soil C were equivalent to 12% of the cumulative annual C fluxes by combustion of plant biomass and 30% of the decrease in the total plant biomass C in a plot. Comparing paired simulations at each site, either including or excluding N losses, illustrated that fire-driven N losses reduced cumulative NPP by about 5% over the entire 63-year period of the simulation on average across sites (Supplementary Fig. 8). The changes in NPP were of substantial magnitude relative to other C fluxes, with the total reduction in C drawdown from NPP being equivalent to 20% of the total annual C emissions from combustion of plant biomass summed over the simulation period, averaged across sites.

We next assessed the potential generality of fire-induced soil C and N losses changing ecosystem C storage and productivity by performing simulations across savanna grasslands globally; these ecosystems represent about 70% of actual global burned area⁷ (see Supplementary Information). When all locations were burned at a biennial frequency, declines in soil C stocks were equivalent to 40% of the changes in

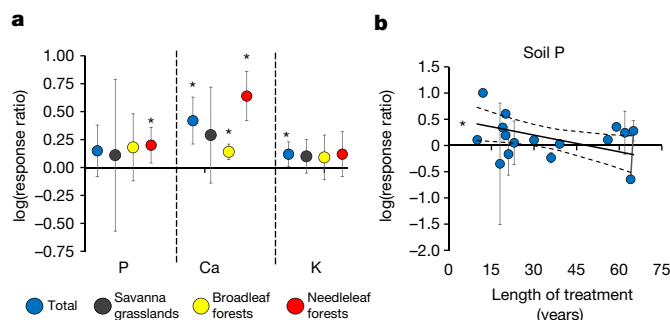


Figure 3 | Responses of P, Ca and K to changes in fire frequency. **a**, Logarithmic response ratios of the concentrations of P ($n = 16$), Ca ($n = 16$) and K ($n = 18$) for the total dataset compiled and partitioned into different ecosystem categories. The response ratio is defined as the concentration of P, Ca or K in elevated plots divided by the concentration in protected plots. **b**, Change in the logarithmic response ratio of soil P as a function of the length of time during which plots experienced contrasting fire frequencies. Error bars in **a** indicate the 95% confidence intervals and those in **b** indicate the variance around the response ratio and dashed lines in **b** are 95% confidence intervals, with an asterisk indicating significant effects ($P < 0.05$). See Supplementary Tables 6 and 7 for statistics.

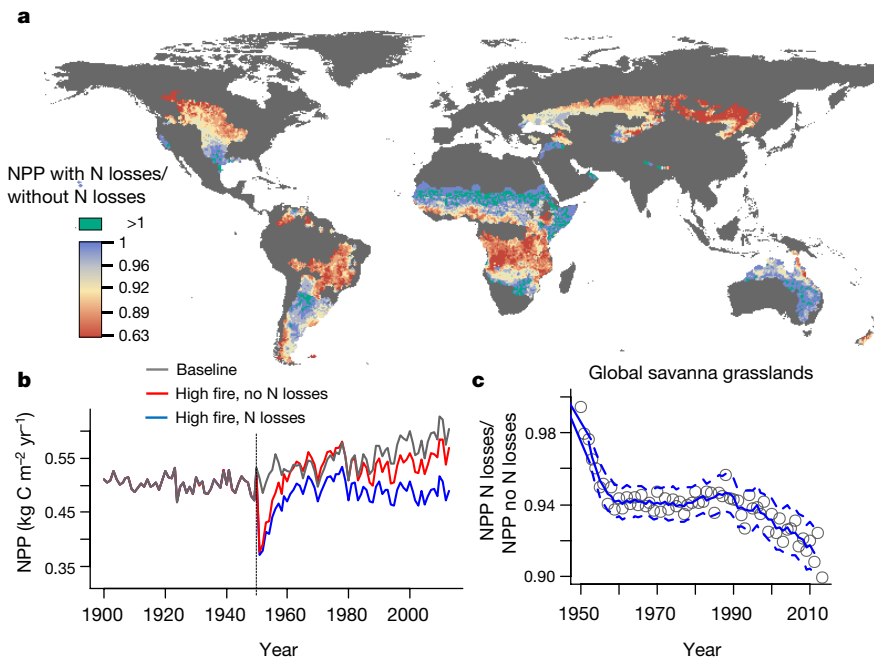


Figure 4 | Effect of N losses on net primary productivity (NPP) across savanna grasslands globally. Simulations were run by initiating a high fire frequency in 1950 (with grid cells burned every two years) and tracking NPP until 2013 with and without N losses. **a**, Relative ratio of cumulative NPP between the two scenarios, with the colour bar scaled by quantiles (values are minimum (0.63), first, second and third quantiles (0.89, 0.92 and 0.96), and maximum (1)). Green cells illustrate areas where N losses stimulated NPP (where the ratio is greater than 1). **b**, Mean NPP simulated across savanna grasslands, weighted by the area of a grid cell. The sharp reduction in NPP in 1950 (grey vertical line) is caused by the initiation of

the prescribed higher fire frequency scenarios, where N is lost by fire (blue line) or not lost by fire (red line). The grey line shows the evolution of NPP as predicted internally in the dynamic global vegetation model LPJ-GUESS with fires determined via BLAZE operating dynamically (for example, as for the period before 1950). **c**, Model simulations of the ratio between NPP with N losses versus without N losses through time, averaged across savanna grasslands globally (each circle is a global average within a year); the solid line represents a five-year rolling average and the dashed lines represent the standard errors across grid cells.

plant biomass C stocks, on average, with the relative contribution of declines in soil C being greatest in driest locations (Supplementary Fig. 11; $r^2 = 0.45$). Furthermore, N losses resulted in widespread declines in NPP (Fig. 4a), with the largest effect on NPP seen in wet tropical regions, probably because of higher potential productivity and N demand. The effect of N losses on NPP increased through time (Fig. 4b, c), amounting to a 9% reduction of NPP in savanna grasslands globally when summed over the entire simulation period and area. Consequently, omitting the multidecadal changes in soil pools that result from shifting fire frequencies may substantially underestimate ecosystem C losses.

Our results reveal several factors that regulate how fire affects C and N in soils, and shed light on potential responses under future fire regimes. First, the effect of fire on both C and N strengthened through time and emerged only over multiple decades. The lack of a saturating response was surprising, and suggests that shifts in fire frequency during the twenty-first century³ may alter soil C and N over an extensive land area. Considering changes in soil C over longer time periods—especially through the formation of pyrogenic C, which can influence long-term C storage and nutrient dynamics^{22,23}—will provide additional insight into the stability of C in the soils and when effects may saturate.

Second, whether fire changed soil C and N and by how much depended on vegetation type across our analysis. The enrichment of N in needleleaf forest soils could be attributable to a number of processes, such as colonization by N-fixing plant species²⁴ or redistribution of mobilized N during the smouldering of the thick forest floor that is characteristic of needleleaf forests²⁵. Whether our results from needleleaf forests that primarily received frequent, low-intensity prescribed fires are representative of colder needleleaf forests that experience less frequent, but more intense, wildfires requires further evaluation, especially for boreal forests. Although we found qualitatively similar responses

of boreal and temperate needleleaf forests, more boreal studies in particular are needed to test the generality in the response and application over longer fire-return intervals and for severe crown fires that can consume the soil organic layer²⁶. Studies of gradients in long-term fire frequencies are lacking at present and do not always examine changes in mineral soils (see, for example, ref. 26).

Further consideration is also needed for relatively wet ecosystems, such as some tropical rainforests, that are now experiencing more frequent burning because of human activities and drying climates²⁷. More frequent slash-and-burn cycles, for example, have been shown to deplete soil C, N and P²⁸ in tropical rainforests. Our observation that the initial P enrichment fades through time may be a critical component in determining the response of P-limited tropical rainforests²⁹ to changes in fire frequency.

Projecting the effect of changes in fire frequency on ecosystem C storage also needs better understanding of historical fire regimes. We compared historical fire frequencies to our elevated and protected fire treatments by using data from a subset of the locations included in the meta-analysis ($n = 25$) that had intermediate fire frequencies to approximate historical natural burning (see Supplementary Information). Compared with these intermediate fire frequencies, more frequent burning significantly decreased C and N concentrations (-13% C and N, $P = 0.007$ and $P < 0.001$, respectively), whereas less frequent burning significantly increased C and N concentrations ($+19\%$ C and $+18\%$ N, $P = 0.0005$ and $P < 0.0001$, respectively) in savanna grasslands (Supplementary Table 8 and Supplementary Fig. 12). Analyses of broadleaf forest sites had less statistical power, but suggested that differences occurred primarily because of greater losses in elevated-frequency relative to historical-frequency plots. In needleleaf forests, fire tended to enrich N in historical-frequency versus protected plots, but elevated versus historical-frequency plots were comparable. Consequently, the significant changes we observed when

comparing elevated-frequency versus protected plots are attributable both to C and N accumulation during fire protection, and to C and N loss during increased burning.

In conclusion, our results reveal the sensitivity of surface soils to fire and the substantial effects that changes in soil pools have on long-term ecosystem C exchange. The large empirical and conservative model-based estimates of soil C changes suggest that present estimates of fire-driven C losses⁷, which primarily consider losses from plant biomass pools, may substantially underestimate the effects of long-term trends in fire frequencies in savanna grasslands and broadleaf forests in particular. Our findings suggest that future alterations in fire regimes in savanna grasslands and broadleaf forests may shift ecosystem C storage by changing soil C levels and changing the N limitation of plant growth, altering the carbon-sink capacity of these fire-prone ecosystems.

Data Availability The datasets generated and analysed during this study are available from the corresponding author on request and in the corresponding papers cited in Supplementary Information.

Received 18 April; accepted 19 October 2017.

Published online 11 December 2017.

- Westerling, A. L., Hidalgo, H. G., Cayan, D. R. & Swetnam, T. W. Warming and earlier spring increase western US forest wildfire activity. *Science* **313**, 940–943 (2006).
- Knorr, W., Arneth, A. & Jiang, L. Demographic controls of future global fire risk. *Nat. Clim. Chang.* **6**, 781–785 (2016).
- Andela, N. *et al.* A human-driven decline in global burned area. *Science* **356**, 1356–1362 (2017).
- Jobbágy, E. G. & Jackson, R. B. The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecol. Appl.* **10**, 423–436 (2000).
- Jackson, R. B., Banner, J. L., Jobbágy, E. G., Pockman, W. T. & Wall, D. H. Ecosystem carbon loss with woody plant invasion of grasslands. *Nature* **418**, 623–626 (2002).
- Randerson, J. T. *et al.* The impact of boreal forest fire on climate warming. *Science* **314**, 1130–1132 (2006).
- van der Werf, G. R. *et al.* Global fire emissions estimates during 1997–2016. *Earth Syst. Sci. Data* **9**, 697–720 (2017).
- LeBauer, D. S. & Treseder, K. K. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology* **89**, 371–379 (2008).
- Pellegrini, A. F. A., Hoffmann, W. A. & Franco, A. C. Carbon accumulation and nitrogen pool recovery during transitions from savanna to forest in central Brazil. *Ecology* **95**, 342–352 (2014).
- DeLuca, T. H. & Sala, A. Frequent fire alters nitrogen transformations in ponderosa pine stands of the inland northwest. *Ecology* **87**, 2511–2522 (2006).
- Reich, P. B., Peterson, D. W., Wedin, D. A. & Wragge, K. Fire and vegetation effects on productivity and nitrogen cycling across a forest-grassland continuum. *Ecology* **82**, 1703–1719 (2001).
- Newland, J. A. & DeLuca, T. H. Influence of fire on native nitrogen-fixing plants and soil nitrogen status in ponderosa pine–Douglas-fir forests in western Montana. *Can. J. For. Res.* **30**, 274–282 (2000).
- Boerner, R. E. J., Huang, J. & Hart, S. C. Impacts of fire and fire surrogate treatments on forest soil properties: a meta-analytical approach. *Ecol. Appl.* **19**, 338–358 (2009).
- Wan, S., Hui, D. & Luo, Y. Fire effects on nitrogen pools and dynamics in terrestrial ecosystems: a meta-analysis. *Ecol. Appl.* **11**, 1349–1365 (2001).
- Neary, D. G., Klopatek, C. C., DeBano, L. F. & Ffolliott, P. F. Fire effects on belowground sustainability: a review and synthesis. *For. Ecol. Manage.* **122**, 51–71 (1999).
- Kauffman, J. B., Cummings, D. L., Ward, D. E. & Babbitt, R. Fire in the Brazilian Amazon: 1. Biomass, nutrient pools, and losses in slashed primary forests. *Oecologia* **104**, 397–408 (1995).
- Pellegrini, A. F. A., Hedin, L. O., Staver, A. C. & Govender, N. Fire alters ecosystem carbon and nutrients but not plant nutrient stoichiometry or composition in tropical savanna. *Ecology* **96**, 1275–1285 (2015).
- Knapp, A. K. & Seastedt, T. R. Detritus accumulation limits productivity of tallgrass prairie. *Bioscience* **36**, 662–668 (1986).
- Hedges, L. V., Gurevitch, J. & Curtis, P. S. The meta-analysis of response ratios in experimental ecology. *Ecology* **80**, 1150–1156 (1999).
- DeBano, L. F., Neary, D. G. & Ffolliott, P. F. *Fire Effects on Ecosystems* (John Wiley, 1998).
- Smith, B. *et al.* Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences* **11**, 2027–2054 (2014).
- Santín, C. *et al.* Towards a global assessment of pyrogenic carbon from vegetation fires. *Glob. Change Biol.* **22**, 76–91 (2016).
- DeLuca, T. H., Aplet, G. H. Charcoal and carbon storage in forest soils of the Rocky Mountain West. *Front. Ecol. Environ.* **6**, 18–24 (2008).
- DeLuca, T. H., Zackrisson, O., Nilsson, M.-C. & Sellstedt, A. Quantifying nitrogen-fixation in feather moss carpets of boreal forests. *Nature* **419**, 917–920 (2002).
- McKee, W. H. *Changes In Soil Fertility Following Prescribed Burning On Coastal Plain Pine Sites* (Southeastern Forest Experiment Station, 1982).
- Wardle, D. A., Hörnberg, G., Zackrisson, O., Kalela-Brundin, M. & Coomes, D. A. Long-term effects of wildfire on ecosystem properties across an island area gradient. *Science* **300**, 972–975 (2003).
- Davidson, E. A. *et al.* The Amazon basin in transition. *Nature* **481**, 321–328 (2012).
- Ramakrishnan, P. S. & Toky, O. P. Soil nutrient status of hill agro-ecosystems and recovery pattern after slash and burn agriculture (Jhum) in north-eastern India. *Plant Soil* **60**, 41–64 (1981).
- Vitousek, P. M., Porder, S., Houlton, B. Z. & Chadwick, O. A. Terrestrial phosphorus limitation: mechanism, implications, and nitrogen-phosphorus interactions. *Ecol. Appl.* **20**, 5–15 (2010).
- Ricklefs, R. E. *The Economy of Nature* (WH Freeman, 2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank all authors of the studies used in the meta-analysis; the Cedar Creek Long Term Ecological Research programme; The Morton Arboretum Center for Tree Science programme; and J. Harden, L. Hedin, S. Pacala and M. Turner for providing feedback. Funding was provided by a National Oceanic and Atmospheric Administration (NOAA) Climate and Global Change Postdoctoral Fellowship (to A.F.A.P.); the Gordon and Betty Moore Foundation (R.B.J.); the Modelling the Regional and Global Earth system (MERGE) (L.P.N.); and the Department of Energy Office of Science Biological and Environmental Research (J.T.R.).

Author Contributions A.F.A.P. and R.B.J. conceived of and designed the study, with input from A.A.; A.F.A.P., S.E.H., P.B.R., B.C.S. and A.J. collected and contributed data; A.F.A.P. performed statistical analyses; L.P.N. developed the fire model; and L.P.N. and A.A. performed model simulations. A.F.A.P. wrote the first draft and all authors contributed feedback.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to A.F.A.P. (afapelle@stanford.edu).

Reviewer Information Nature thanks T. DeLuca, A. D. McGuire and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Successful conservation of global waterbird populations depends on effective governance

Tatsuya Amano^{1,2}, Tamás Székely^{3,4}, Brody Sandel⁵, Szabolcs Nagy⁶, Taej Mundkur⁶, Tom Langendoen⁶, Daniel Blanco⁷, Candan U. Soykan⁸ & William J. Sutherland¹

Understanding global patterns of biodiversity change is crucial for conservation research, policies and practices. However, for most ecosystems, the lack of systematically collected data at a global level limits our understanding of biodiversity changes and their local-scale drivers. Here we address this challenge by focusing on wetlands, which are among the most biodiverse and productive of any environments^{1,2} and which provide essential ecosystem services^{3,4}, but are also amongst the most seriously threatened ecosystems^{3,5}. Using birds as an indicator taxon of wetland biodiversity, we model time-series abundance data for 461 waterbird species at 25,769 survey sites across the globe. We show that the strongest predictor of changes in waterbird abundance, and of conservation efforts having beneficial effects, is the effective governance of a country. In areas in which governance is on average less effective, such as western and central Asia, sub-Saharan Africa and South America, waterbird declines are particularly pronounced; a higher protected area coverage of wetland environments facilitates waterbird increases, but only in countries with more effective governance. Our findings highlight that sociopolitical instability can lead to biodiversity loss and undermine the benefit of existing conservation efforts, such as the expansion of protected area coverage. Furthermore, data deficiencies in areas with less effective governance could lead to underestimations of the extent of the current biodiversity crisis.

Quantifying global patterns of biodiversity change is essential for assessing anthropogenic impacts on biodiversity, conservation priorities and the effectiveness of conservation efforts^{6,7}. It has therefore been identified as a research priority by major international bodies^{8,9}. However, most taxa have serious gaps in the spatial extent and resolution covered by available biodiversity data¹⁰, and our current view of global biodiversity change is therefore limited to coarse-resolution patterns¹¹, data-rich countries¹² or protected areas¹³. This has impeded the identification of hotspots of abundance loss, and the analysis of the effects of local-scale drivers on biodiversity change at the global scale (see Supplementary Discussion; also see Supplementary Information for the Abstract in different languages).

Globally, wetlands cover more than 1,280 million hectares of coastal, inland and human-made habitats^{3,14}. Despite their high levels of biological diversity and productivity^{1,2} and the crucial ecosystem functions and services they provide^{1,3,4}, wetlands have been degraded and lost at higher rates than any other ecosystem³. However, the lack of appropriate data has hampered assessments of changes in wetland biodiversity at a global scale.

Here we address this by examining waterbirds as an indicator taxon for assessing the status of biodiversity in wetland ecosystems. Waterbirds have a long history of systematic monitoring, and therefore present a global dataset of abundance changes with unusually high

spatial extent and resolution¹⁵. Modelling the global data for waterbirds enabled us to test two fundamental questions that are rarely explored in tandem; we asked where global changes in species abundance have been concentrated and what might explain changes in abundance at community, species and population levels. For the second question, we tested hypothesized predictors that were categorized into three groups: (i) anthropogenic effects (surface water change, economic and human population growth, agricultural expansion and climate change), (ii) conservation efforts and effectiveness (protected area coverage and governance), and (iii) biological characteristics of species (range size, migratory status and body size) (Extended Data Table 1). Our dataset comprised 2,463,403 count records, covering the months of January–February for the past three decades and recording 461 waterbird species at 25,769 survey sites throughout the globe (Extended Data Fig. 1). Using a hierarchical Bayesian model, we estimated the global distribution of changes in the abundance of each species between 1990 and 2013 at $1^\circ \times 1^\circ$ spatial resolution (Supplementary Data 1). We then summarized the changes at three levels: mean changes in abundance across all waterbird species present in each grid cell (community-level changes), mean changes across all grid cells for each species (species-level changes) and changes in each grid cell for each species (population-level changes).

In most species, population-level changes in abundance varied markedly across geographical ranges. Some species that have increased in abundance in Europe showed severe declines in other regions (Fig. 1a–c) and vice versa (see Supplementary Data 1). Declines were especially pronounced in Africa for grebes, flamingos, pelicans, cormorants and shorebirds, in South America for shorebirds, storks, ibises, herons, waterfowl, cranes and rails, and in western and central Asia for waterfowl, cranes and rails (Fig. 1d–k).

We found major community-level abundance losses in areas in which biodiversity assessments have been limited, namely western and central Asia, sub-Saharan Africa and South America (Fig. 2a). On average, community-level declines were most severe in South America, which has experienced a 0.95% annual decline that equates to a 21% total decline over 25 years (Fig. 2b). The declines were also severe in western and central Asia, but predominantly occurred inland rather than in coastal regions. By contrast, Europe has experienced community-level increases in waterbird abundance, though even in regions that experienced these increases some species showed severe abundance declines (Supplementary Data 1). These geographic patterns predominantly reflected patterns in migrant species (Extended Data Fig. 2a), as non-migrants were observed only in some regions; non-migrants showed community-level declines in South America and parts of east Asia, south Asia and southeast Asia (Extended Data Fig. 2b).

Of the eight explanatory variables representing anthropogenic impacts and conservation efforts and effectiveness (see Methods),

¹Conservation Science Group, Department of Zoology, University of Cambridge, The David Attenborough Building, Pembroke Street, Cambridge, CB2 3QZ, UK. ²Centre for the Study of Existential Risk, University of Cambridge, 16 Mill Lane, Cambridge, CB2 1SG, UK. ³Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Bath, BA2 7AY, UK. ⁴Department of Evolutionary Zoology, University of Debrecen, Debrecen, H-4010, Hungary. ⁵Department of Biology, Santa Clara University, 500 El Camino Real, Santa Clara, California 95053, USA. ⁶Wetlands International Head Office, Horapark 9, 6717 LZ Ede, The Netherlands. ⁷Wetlands International LAC Argentina Office, Capitán General Ramón Freire 1512, Buenos Aires 1426, Argentina. ⁸National Audubon Society, Conservation Science, 220 Montgomery St., Suite 1000, San Francisco, California 94104, USA.

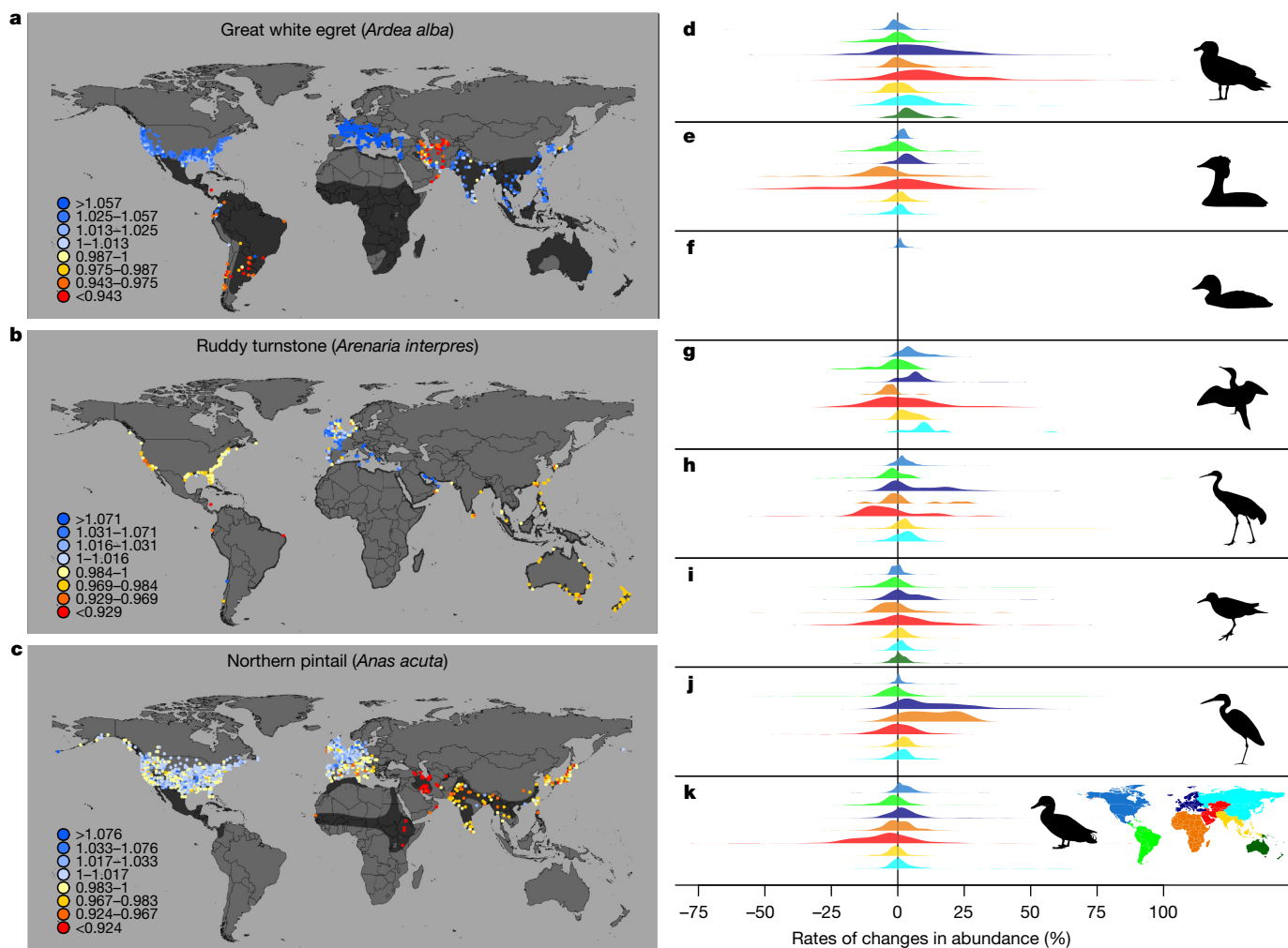


Figure 1 | Population-level changes in waterbird abundance in each $1^\circ \times 1^\circ$ grid cell between 1990 and 2013. a–c, Examples of population-level abundance changes, for *Ardea alba* (a), *Arenaria interpres* (b) and *Anas acuta* (c). Red, declines; blue, increases; dark grey shading, non-breeding geographical range of the species. d–k, Histograms of population-level changes for all species in each of the eight taxa, at all grid cells in each region shown in the inserted map (see Methods for the definition of each species group). Silhouettes reproduced from PhyloPic (<http://phylopic.org/>) under a Creative Commons licence

(<http://creativecommons.org/licenses/by/3.0/>) (d–g, i–k) or Public Domain Dedication licence (<http://creativecommons.org/publicdomain/zero/1.0/>) (h). d, i, Rebecca. Groom; e, f, Doug Backlund (photo) (e) or Unknown (photo) (f), John E. McCormack, Michael. G. Harvey, Brant. C. Faircloth, Nicholas. G. Crawford, Travis. C. Glenn, Robb. T. Brumfield & T. Michael. Keesay; g, j, Shyamal/Wikimedia Commons; k, Maija. Karala (image flipped horizontally). Map produced from Natural Earth data v.1.4.0 (<http://www.naturalearthdata.com/>).

governance—defined as how effectively the authorities of a country exercise rules and enforcement mechanisms—was the strongest predictor of community-level abundance changes (Fig. 3a). Waterbird communities experienced the greatest declines in countries with less effective governance (for example, countries in western and central Asia or South America), and increased in countries in which governance was

more effective (for example, countries in Europe and North America, Fig. 3b). The effects of governance also interacted with those of protected area coverage (Fig. 3a); it was only in areas with more effective governance that extensive protected area coverage was associated with community-level increases in waterbird abundances (Extended Data Fig. 3a). Community-level declines were also pronounced in areas with

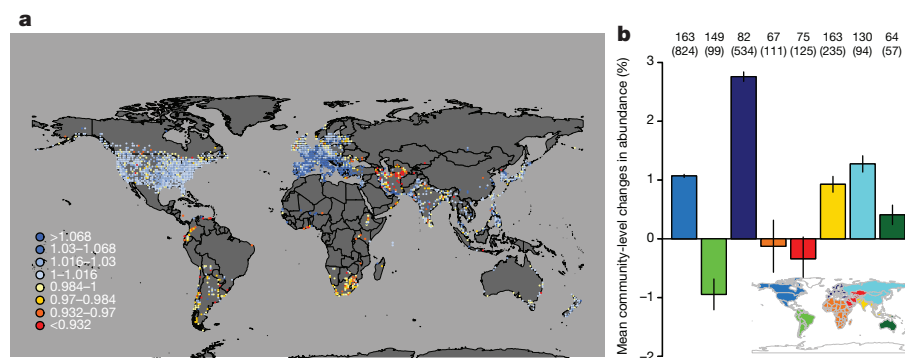


Figure 2 | Mean changes in abundance across 461 waterbird species (community-level changes) between 1990 and 2013. a, b, Global distribution (a) and mean with 95% confidence intervals (b) across all grid cells in each region shown in the inserted map. Numbers above bars, numbers of species observed; numbers of grid cells in parentheses. Map produced from Natural Earth data v.1.4.0 (<http://www.naturalearthdata.com/>).

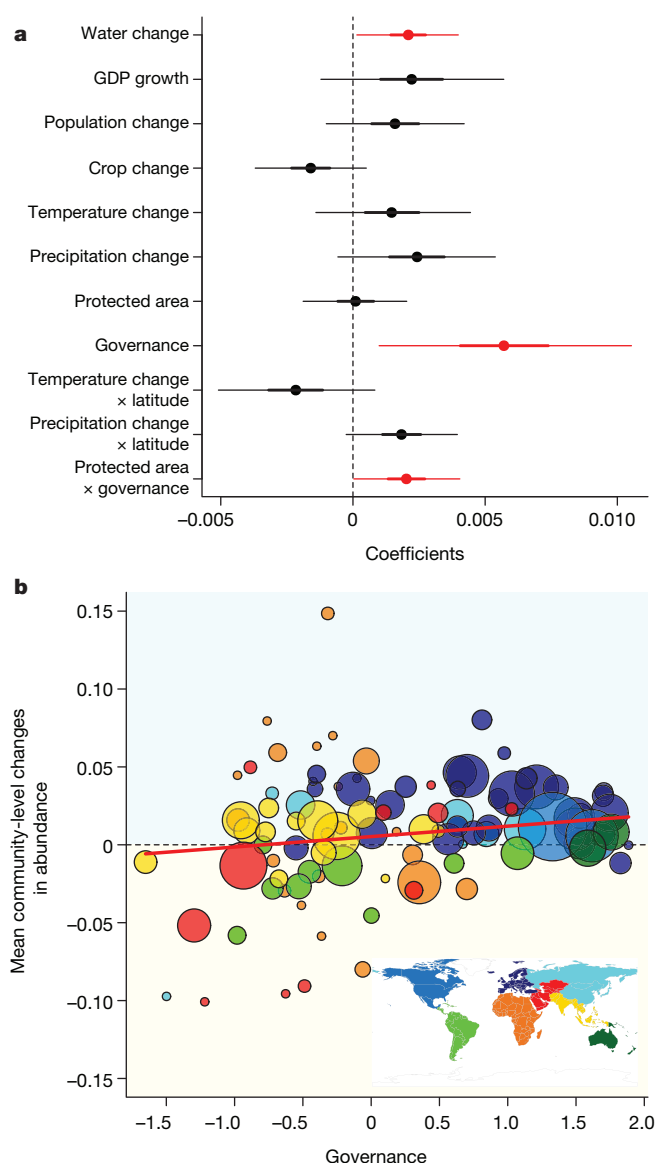


Figure 3 | Effects of predictors on community-level changes in waterbird abundance. **a**, Estimated coefficients in the multivariate analysis ($n = 2,079$). Posterior medians with 95% and 50% (thick lines) credible intervals are shown. Coefficients with 95% credible intervals that do not overlap with zero are shown in red. The coefficients represent the effect size of the standardized variables. **b**, Relationship between community-level changes and countries' governance. Each circle represents a country; circle size, the number of $1^\circ \times 1^\circ$ grid cells with estimates; colour indicates the region shown in the inset map; regression line shown in red. Map produced from Natural Earth data v.1.4.0 (<http://www.naturalearthdata.com/>).

a higher rate of surface water loss (for example, western and central Asia¹⁶, Extended Data Fig. 3b).

To explore the possible causes of community-level changes, we partitioned the effects of explanatory variables into species-level (explaining variations in species-level changes between species) and population-level effects (explaining variations in population-level changes within species) for 293 species with sufficient data. Species-level changes were explained by the interaction between governance and protected area coverage, by gross domestic product (GDP) growth rates and by body mass (Fig. 4a). Consistent with the community-level analysis, waterbird species with a higher coverage of protected areas increased more, but only in countries with more effective governance (Fig. 4c). Species in countries with rapidly growing economies, as well

as small-bodied species, experienced greater declines (Fig. 4b, d). Governance was also the best predictor of population-level abundance changes, and most of the species that were significantly affected by governance showed larger population-level declines in areas with less effective governance (Extended Data Fig. 4 and Supplementary Discussion). These conclusions were robust even when considering the correlation between governance and GDP per capita, and were also robust to other sensitivity analyses (Extended Data Figs 5–7, Supplementary Discussion).

Although our data are not spatially complete (Extended Data Fig. 1 and Supplementary Discussion), by quantifying abundance changes within each species over large geographic areas we uncovered new hotspots of threats to bird species in wetland ecosystems. Previous studies (see Supplementary Discussion) did not identify biodiversity loss in, for example, western and central Asia, mainly because relevant data were unavailable. This spatial overlap between general data gaps and biodiversity loss could cause an underestimation of the ongoing biodiversity crisis, which highlights the need for global monitoring of species' abundances.

Our results emphasize the importance of governance—presumably the environmental aspects of governance (see Methods)—in explaining global patterns in waterbird abundance changes. Local and regional studies have increasingly highlighted the environmental consequences of ineffective governance, such as species population declines¹⁷, deforestation¹⁸ and agricultural expansion¹⁹. Ineffective governance is often associated with the absence of positive attitudes to environmental protection, weakly enforced environmental legislation and low levels of investment in conservation^{20–22}, leading to habitat loss and degradation. For example, unsustainable water management and dam construction in western and central Asia have caused drastic losses in permanent water over the past 30 years¹⁶. As a result, in Iran even some wetlands designated as protected areas have dried out²³. In South America, wetlands in central Argentina lack legal protection or regulations on water use, and many have been lost²⁴. Ineffective hunting regulations can also explain decreases in abundance under conditions of ineffective governance. Political instability can weaken the legal enforcement of hunting regulations and thereby promote unsustainable and often illegal killing, even in protected areas²⁵; numerous waterbird species are under severe hunting pressure in Iran²³ and South America²⁶. As wetland loss and hunting pressure are the main threats to most taxa, the hotspots of waterbird declines identified here merit urgent attention as areas of potential loss and degradation of wetland biodiversity, and its concomitant functions and services.

This study corroborates the observation that protected areas improve the conservation status of waterbird species, although the benefits of these protected areas are applicable only in countries with more effective governance. Our results provide strong support at the global scale for the argument that effective governance is critical for protected areas to achieve their goals²⁷. Even in developing countries with less effective governance, protected area coverage can be high (Extended Data Fig. 8); however, these protected areas have been insufficient to maintain stable waterbird populations since 1990. By contrast, in wealthier regions with more effective governance, such as Western Europe, waterbirds have responded positively to the establishment of refuges and stronger legal protection under measures governed by the EU Birds Directive²⁸.

Although the global coverage of protected areas continues to increase, our findings indicate that ineffective governance could undermine the benefits of such conservation efforts that aim to improve the status of global biodiversity. Levels of governance should be considered in the processes of identifying and prioritising areas of conservation importance, and distributing future research and funding efforts. There is also an urgent need to measure, monitor, improve and raise awareness about environmental governance globally. Global conservation conventions and specific agreements and frameworks could mobilize international resources and expertise to strengthen effective

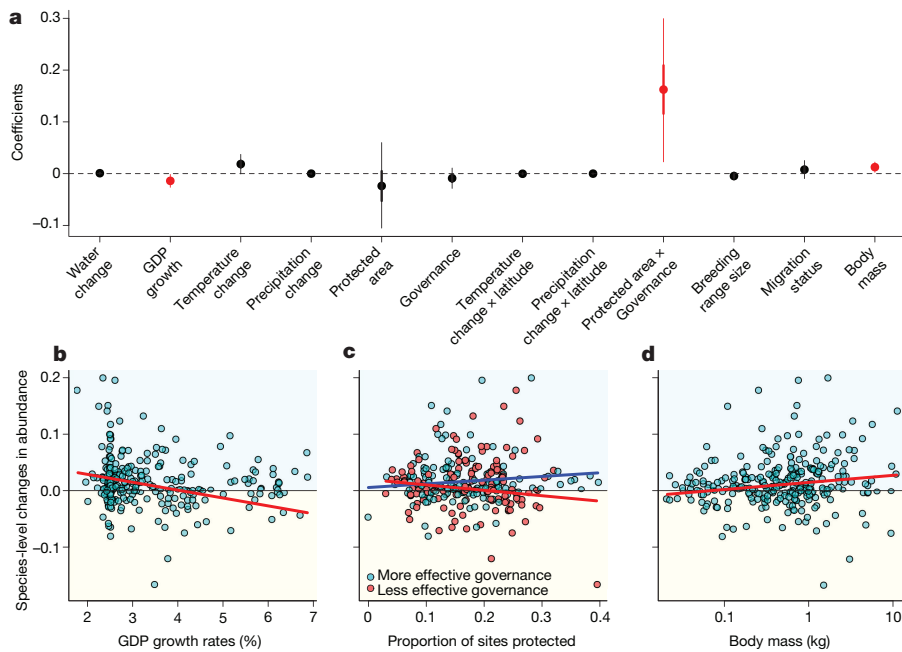


Figure 4 | Effects of predictors on species-level abundance changes in 293 waterbird species that were recorded in at least ten grid cells. **a**, Estimated coefficients with 95% and 50% (thick lines) credible intervals; those not overlapping with zero are shown in red. **b–d**, Relationship at the species level between abundance changes and GDP growth rates (**b**), proportion of sites covered by protected areas (**c**) and body mass (**d**). Values and regression lines for species in areas with more (above median) and less (below median) effective governance are shown in blue and red, respectively, in **c**. See Supplementary Data 2 for details of the 293 species.

governance. Governance is now recognized to be essential for economic growth, social development and the eradication of poverty and hunger⁴. Efforts to better understand and improve governance, as well as to find means of improving the effectiveness of specific measures when governance is weak, therefore provide common ground for conservationists, social scientists, policy makers and the public for achieving sustainable development.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 May; accepted 16 November 2017.

Published online 20 December 2017.

1. The Ramsar Convention on Wetlands. *Wetland Ecosystem Services*. [\(http://archive.ramsar.org/cda/en/ramsar-pubs-info-ecosystem-services/main/ramsar/1-30-103%5E24258_4000_0_\)](http://archive.ramsar.org/cda/en/ramsar-pubs-info-ecosystem-services/main/ramsar/1-30-103%5E24258_4000_0_) (2011).
2. Dudgeon, D. *et al.* Freshwater biodiversity: importance, threats, status and conservation challenges. *Biol. Rev. Camb. Philos. Soc.* **81**, 163–182 (2006).
3. Millennium Ecosystem Assessment. *Ecosystems and Human Well-Being: Wetlands and Water Synthesis* (World Resources Institute, 2005).
4. United Nations General Assembly. *Transforming Our World: the 2030 Agenda for Sustainable Development. Resolution Adopted by the General Assembly on 25 September 2015* (United Nations, 2015).
5. Young, H. S., McCauley, D. J., Galetti, M. & Dirzo, R. Patterns, causes, and consequences of Anthropocene defauna. *Annu. Rev. Ecol. Evol. Syst.* **47**, 333–358 (2016).
6. Balmford, A., Green, R. E. & Jenkins, M. Measuring the changing state of nature. *Trends Ecol. Evol.* **18**, 326–330 (2003).
7. Margules, C. R. & Pressey, R. L. Systematic conservation planning. *Nature* **405**, 243–253 (2000).
8. Convention on Biological Diversity. *Decision X/2. The Strategic Plan for Biodiversity 2011–2020 and the Aichi Biodiversity Targets* (Secretariat of the Convention on Biological Diversity, 2010).
9. Intergovernmental Platform on Biodiversity and Ecosystem Services. *Generic Scoping Report for the Regional and Subregional Assessments of Biodiversity and Ecosystem Services* (Intergovernmental Platform on Biodiversity and Ecosystem Services, 2015).
10. Pimm, S. L. *et al.* The biodiversity of species and their rates of extinction, distribution, and protection. *Science* **344**, 1246752 (2014).
11. WWF. *Living Planet Report 2016. Risk and Resilience in a New Era* (WWF International, 2016).
12. Bowler, D. E. *et al.* Cross-realm assessment of climate change impacts on species' abundance trends. *Nat. Ecol. Evol.* **1**, 0067 (2017).
13. Barnes, M. D. *et al.* Wildlife population trends in protected areas predicted by national socio-economic metrics and body size. *Nat. Commun.* **7**, 12747 (2016).
14. The Ramsar Convention on Wetlands. *Classification System for Wetland Type* http://archive.ramsar.org/cda/en/ramsar-documents-guidelines-strategic-framework-and/main/ramsar/1-31-105%5E20823_4000_0_#B (2012).
15. Boere, G. C., Galbraith, C. A. & Stroud, D. A. eds. *Waterbirds Around the World* (The Stationery Office, 2006).

16. Pekel, J.-F., Cottam, A., Gorelick, N. & Belward, A. S. High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418–422 (2016).
17. Smith, R. J., Muir, R. D. J., Walpole, M. J., Balmford, A. & Leader-Williams, N. Governance and the loss of biodiversity. *Nature* **426**, 67–70 (2003).
18. Urnemiya, C., Rametsteiner, E. & Kraxner, F. Quantifying the impacts of the quality of governance on deforestation. *Environ. Sci. Policy* **13**, 695–701 (2010).
19. Ceddia, M. G., Bardsley, N. O., Gomez-y-Paloma, S. & Sedlacek, S. Governance, agricultural intensification, and land sparing in tropical South America. *Proc. Natl Acad. Sci. USA* **111**, 7242–7247 (2014).
20. Harring, N. Understanding the effects of corruption and political trust on willingness to make economic sacrifices for environmental protection in a cross-national perspective. *Soc. Sci. Q.* **94**, 660–671 (2013).
21. Sundström, A. Covenants with broken swords: corruption and law enforcement in governance of the commons. *Glob. Environ. Change* **31**, 253–262 (2015).
22. Miller, D. C., Agrawal, A. & Timmons Roberts, J. Biodiversity, governance, and the allocation of international aid for conservation. *Conserv. Lett.* **6**, 12–20 (2013).
23. Nourani, E., Kaboli, M. & Collen, B. An assessment of threats to Anatidae in Iran. *Bird Conserv. Int.* **25**, 242–257 (2015).
24. Brandolin, P. G. & Blendinger, P. G. Effect of habitat and landscape structure on waterbird abundance in wetlands of central Argentina. *Wetl. Ecol. Manag.* **24**, 93–105 (2016).
25. Brochet, A.-L. *et al.* Preliminary assessment of the scope and scale of illegal killing and taking of birds in the Mediterranean. *Bird Conserv. Int.* **26**, 1–28 (2016).
26. Morrison, R. I. G. *et al.* Dramatic declines of semipalmated sandpipers on their major wintering areas in the Guianas, Northern South America. *Waterbirds* **35**, 120–134 (2012).
27. Lockwood, M. Good governance for terrestrial protected areas: A framework, principles and performance outcomes. *J. Environ. Manage.* **91**, 754–766 (2010).
28. Kirby, J. S. *et al.* Key conservation issues for migratory land- and waterbird species on the world's major flyways. *Bird Conserv. Int.* **18**, S49–S73 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the coordinators, thousands of volunteer counters and funders of the International Waterbird Census and Christmas Bird Count (see Supplementary Notes for information on funders); D. Unterkofler for preparing the NWC data, H. Okamura for statistical advice, J. P. González-Varo for his comments on an earlier draft and M. Amano for long-standing support.

Author Contributions T.A., T.S. and W.J.S. designed the study. T.A., T.S., B.S., S.N., T.M., T.L., D.B. and C.U.S. collected and prepared data for the analyses. T.A. analysed the data and wrote the paper. All authors discussed the results and commented on the manuscript at all stages.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to T.A. (amatatsu830@gmail.com).

Reviewer Information *Nature* thanks R. Fuller and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Data. *Waterbird count data.* Data used in this study consisted of site-specific annual counts from the International Waterbird Census (IWC) coordinated by Wetlands International²⁹ and the Christmas Bird Count (CBC) coordinated by the National Audubon Society³⁰.

Launched in 1967, the IWC is a scheme involving more than 15,000 observers that monitors waterbird numbers and covers more than 25,000 sites in over 100 countries. The IWC is divided into four regions, each of which corresponds to a major migratory flyway of the world: the African–Eurasian Waterbird Census (AEWC), Asian Waterbird Census (AWC), Caribbean Waterbird Census (CWC) and Neotropical Waterbird Census (NWC). We did not use data from the CWC, because it started only in 2010 and therefore provides only short-term data. The survey methodology is essentially the same across the four regional schemes. Population counts are typically carried out once every year in mid-January. Additional counts are conducted in other months, particularly in July in the Southern Hemisphere; for consistency, we used only counts from January and February. Our Northern Hemisphere data therefore relate to non-breeding populations, whereas those from the Southern Hemisphere also include some breeding populations. In each country that is covered by the survey, national coordinators manage an inventory of wetland sites (hereafter, survey sites) that include sites of international- or national-level recognition (for example, Ramsar sites, Important Bird Areas, national parks and so on). Each survey site is generally defined by boundaries so that observers know precisely which areas are to be covered in the surveys. The observers consist of a wide variety of volunteers, but national coordinators usually train them using materials produced by Wetlands International to ensure the quality of count data. Survey sites (normally up to a few km²) are typically surveyed by about two observers for up to four hours, but larger sites can require a group of observers to work over several days. The time of survey on any given day depends on the type of survey sites: inland sites are normally surveyed during the morning or late afternoon, whereas coastal sites are surveyed during high tide periods (mangrove areas and nearby mudflats are, however, surveyed during low tides). Surveys cover waterbirds, which are defined as bird species that are ecologically dependent on wetlands²⁹. Counts are usually made by scanning flocks of waterbirds with a telescope or binoculars and counting each species. Zero counts are not always recorded and are thus inferred using a set of criteria (see below). Count records and associated information are submitted to the national coordinators, who compile the submitted records, check their validity and submit them to Wetlands International. Further details of survey methodology have been previously published^{29,31}.

As the IWC does not cover North America, we also used data from the CBC, which has been conducted annually since 1900, involves more than 70,000 observers each year and now includes over 2,400 count circles (defined as survey sites in this study)³². Each CBC consists of a tally of all bird species detected within a survey site (a circle 24.1 km in diameter), on a single day that falls on a date between 14th December and 5th January. The majority of circles (and most historical data) are from the US and Canada. Observers join groups that survey subunits of the circle during the course of the day; they use a variety of transportation methods, mostly surveying on foot or in a car but also using boats, skis, or snowmobiles. The number of observers and the duration of counts vary among circles and through time. The total number of survey hours per count has been recorded as a covariate to account for the variable duration of and participation in the count. In this paper, we only used records describing waterbird species.

We compiled data from each scheme by species, except for data derived from the AEWC that had already been stored by flyway for each species³³. Because data from the NWC are only available after 1990, we restricted the study to data that post-dated 1990 for all regions. The latest records were in 2013. Although the data included 487 waterbird species in total, we excluded from the analyses species with 20 or fewer records; this resulted in 461 species being analysed (see Supplementary Data 2 for the full list of species). For the IWC data, we generated zero counts using an established approach³³. In this approach, we first established a list of all species observed in each country, and assumed a zero count for any species that was on the list but not recorded at a particular site on a particular day (if the site was surveyed on that day), as shown by the presence of any other species' record(s), and if no multi-species code related to the species (for example, *Anatinae* spp. for species of the genus *Anas*) was recorded for the site–date combination. We projected all survey sites onto a Behrmann equal-area cylindrical projection and assigned them to grid cells with a grain size of 96.49 km, or approximately 1° at 30° N or S.

When visualizing the estimated abundance changes (for example, see Figs 2b, 3b), the North and South American regions correspond to regions covered by the CBC and NWC, respectively. The regions covered by the AEWC and AWC were divided into a total of six regions on the basis of socio-economic and ecological differences. The AEWC was divided into three regions: Europe, Africa, and western and central

Asia. The AWC was also divided into three regions: south and southeast Asia, east Asia and Russia, and Oceania.

Explanatory variables. To explain variations in waterbird abundance changes over space and species, we first set up multiple hypotheses on the basis of earlier studies and then identified explanatory variables that represented these hypotheses (Extended Data Table 1). We aggregated all the explanatory variables, except those relating to species characteristics, to the same 1° × 1° grid cells.

As measures of governance we used the Worldwide Governance Indicators, which summarize six dimensions of governance: voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law, and control of corruption³⁴. A previous study¹⁹ of six South American countries found that pro-environmental behaviours are associated with environmental aspects of governance rather than the conventional dimensions of governance represented by the Worldwide Governance Indicators. At the global scale, however, the mean of the Worldwide Governance Indicators was strongly correlated with the Environmental Performance Index (EPI)³⁵, one of the indicators of environmental governance used in the aforementioned study¹⁹ ($r = 0.71$, $n = 180$). This indicates that the Worldwide Governance Indicators are also a good predictor of environmental aspects of governance at the global scale. Further, the EPI consists of multiple indicators, some of which are directly related to our measures of conservation efforts, such as terrestrial protected areas and species protection. We thus decided not to use the EPI in our analysis, as using it together with the coverage of protected areas in our analysis could result in redundancies.

In the World Database on Protected Areas (<https://www.protectedplanet.net/>), not every protected area has information on the year of designation. We therefore calculated the proportion of sites located within any protected area, assuming that this reflects the proportion of sites covered by protected areas designated at least before 2013 (the latest survey year of count data used in this study). To examine the sensitivity of our conclusions to this assumption, we also calculated as the most conservative approach only the proportion of sites covered by protected areas that are known to have been designated before 1990 (the oldest survey year), and conducted the same analyses using this variable (results in Extended Data Fig. 5 and Supplementary Discussion). When assessing the effectiveness of protected areas, confounding factors can mask or mimic the effects of protected areas. We controlled for effects of potential drivers of abundance changes (listed in Extended Data Table 1) by including them together with protected area coverage in the same multivariate models.

On the basis of information from the Birdlife Data Zone (<http://datazone.birdlife.org/home>), the migratory status of the 461 species analysed in this study falls into four categories: full migrant, altitudinal migrant, nomadic and not a migrant. In this study, we defined species that were categorised as full migrant or altitudinal migrant as migrants.

Other data. We derived information on generation length (in years) from the BirdLife Data Zone, and the Red List category assessed by the International Union for Conservation of Nature from the BirdLife Checklist of the Birds of the World³⁶, for each species. Generation length was not available for five species, for which we used the mean values across all species in the same genus. We used generation length as well as the bird species global distribution maps³⁷ for the visualization of results (see Supplementary Data 1 for more detail). Species groups used in Fig. 1 are based on the International Ornithological Congress World Bird List³⁸: coursers, gulls, terns and auks (*Alcidae*, *Glareolidae*, *Laridae* and *Stercorariidae*), grebes and flamingos (*Phoenicopteridae* and *Podicipedidae*), loons and petrels (*Gaviidae* and *Procellariidae*), pelicans, boobies and cormorants (*AnHINGidae*, *Fregatidae*, *Pelecanidae*, *Phalacrocoracidae* and *Sulidae*), rails and cranes (*Aramidae*, *Gruidae* and *Rallidae*), shorebirds (*Burhinidae*, *Charadriidae*, *Dromadidae*, *Haematopodidae*, *Ibidorhynchidae*, *Jacaniidae*, *Recurvirostridae*, *Rostratulidae* and *Scolopacidae*), storks, ibises and herons (*Ardeidae*, *Ciconiidae* and *Threskiornithidae*), and waterfowl (*Anatidae* and *Anhimidae*).

Statistical analyses. *Model for quantifying abundance changes.* To account for missing values, large observation errors and spatial structure in the data, we used a hierarchical Bayesian spatial model and quantified population-level changes in the abundance of each species within each 1° × 1° grid cell. This model is an extension of a model developed and used to quantify waterbird abundance changes in previous studies^{39,40}; it is based on the site effect for site i , overall year effect for year t and the cell-specific year effect for grid cell j and year t . The overall year effect β_t is assumed to be affected by the year effect in the previous two years:

$$\beta_t \sim \text{normal}(\beta_{t-1} + r(\beta_{t-1} - \beta_{t-2}), \sigma_o^2)$$

Here σ_o^2 is the variance of the overall year effect, and r ranges from 0 to 1 and determines the smoothness of the estimated curve. With $r = 0$, the overall year effect is modelled as a simple random-walk process, whereas other values lead to a correlated random walk with different degrees of smoothness (a larger r causes

a more smoothed curve). The cell-specific year effect $\beta_{j(i),t}$ is drawn from a normal distribution with mean β_t as follows:

$$\beta_{j(i),t} \sim \text{normal}(\beta_t, \sigma_\beta^2)$$

Including the variance in the year effect σ_β^2 enables the model to account for variations in trends of population counts among grid cells. The variable $j(i)$ indicates that grid cell j includes site i . Assuming the same population trend across all sites within each grid cell, the mean count $\mu_{i,t}$ at site i in grid cell j and year t is modelled with the cell-specific year effect $\beta_{j(i),t}$, the site effect α_i , the spatially correlated random effect $\gamma_{j(i)}$ and the overdispersion effect $\delta_{i,t}$:

$$\log(\mu_{i,t}) = \alpha_i + \beta_{j(i),t} + \gamma_{j(i)} + \delta_{i,t} \quad (1)$$

Here, α_i and $\delta_{i,t}$ are drawn from a mean-zero normal distribution with variance σ_α^2 and σ_δ^2 , respectively. The variable $\gamma_{j(i)}$ is drawn from an intrinsic Gaussian conditional autoregressive (CAR) prior distribution:

$$\gamma_{j(i)} | \gamma_k \sim \text{normal} \left(\frac{\sum_{j \neq k} w_{j,k} \gamma_k}{n_j}, \frac{\sigma_\gamma^2}{n_j} \right) \quad (2)$$

where $w_{j,k} = 1$ if grid cells j and k are neighbours, and 0 otherwise. The variable n_j is the total number of neighbours of grid cell j ; neighbours are grid cells directly adjacent to grid cell j , and include cells that are diagonally adjacent. The amount of variation between the random effects is controlled by σ_γ^2 . The observed count $y_{i,t}$ in site i and year t is assumed to derive from a Poisson distribution with mean $\mu_{i,t}$.

We assumed constant survey efforts over time for the IWC, because regular and standardized surveys with constant methods, efforts and timing are strongly encouraged in this scheme³¹ (see Supplementary Discussion). However, survey efforts in the CBC are known to vary through time. By using the total number of survey hours per count as the measure of survey efforts, we explicitly accounted for the effort effect for the CBC data following a previously published analysis⁴¹:

$$\log(\mu_{i,t}) = \alpha_i + \beta_{j(i),t} + \gamma_{j(i)} + \delta_{i,t} + \frac{B \left(\left(\frac{\zeta_{i,t}}{\bar{\zeta}} \right)^p - 1 \right)}{p} \quad (3)$$

Here $\zeta_{i,t}$ is the total number of survey hours per count and $\bar{\zeta}$ is the mean value of $\zeta_{i,t}$. The parameters B and p determine a range of relationships between effort and the number of birds counted⁴¹. To test whether accounting for survey efforts changes the conclusions of this paper, we also applied the model without the effort effect to the CBC data, and compared the two models in terms of their estimated rate of abundance change within each grid cell for each of the 159 species with more than two grid cells. The estimated spatial patterns in abundance changes in each of the two models were highly correlated (median Pearson's $r = 0.99$, minimum $r = 0.88$), which indicates that the model without the effort effect that was used for the IWC data is valid. Further discussions on the potential effects of temporal changes in survey efforts are provided in the Supplementary Discussion.

We applied the models to count data for each species at a regional population level. For example, count data for the Eurasian wigeon *Mareca penelope* are separately compiled as five populations: three (northwest European, Black Sea–Mediterranean and southwest Asian–northeast African) in the AEWC, one in the AWC and one in the CBC. In this case, we applied the models separately to each of the five populations. As the result, we analysed 775 regional populations of 461 species (see Supplementary Data 2 for the full list of species). For 38 regional populations in which no grid cells with count records were adjacent to one other, we dropped the spatially correlated random effect $\gamma_{j(i)}$ from equations (1) and (3). For 32 regional populations with only one grid cell that included more than one survey site, we dropped $\gamma_{j(i)}$ and also replaced the cell-specific year effect $\beta_{j(i),t}$ with the overall year effect β_t . For 22 regional populations with only one survey site, we applied a generalized linear model with a Poisson distribution, using observed counts as the response variable and years as the explanatory variable, and used the estimated slope as the rate of abundance change.

Using only grid cells that had on average four or more non-zero records per site, we fitted the models to the data with the Markov chain Monte Carlo (MCMC) method in WinBUGS v.1.4.3⁴² and the R2WinBUGS package⁴³ in R v.3.3.2⁴⁴. Prior distributions of parameters were set as non-informatively as possible, to produce estimates similar to those generated by a maximum likelihood method. We used gamma distributions with a mean of 1 and variance of 100 for the inverses of σ_α^2 , σ_β^2 , σ_δ^2 and σ_γ^2 , normal distributions with a mean of 0 and variance of 100 for β_1 , β_2 and B , a beta distribution with a mean of 0.5 and variance of 0.083 ($\alpha = \beta = 1$), which is a uniform distribution, for r , and a uniform distribution on the interval $[-4, 4]$ for p following a previous study⁴⁵. Each MCMC algorithm was

initially run with three chains with different initial values for 300,000 iterations with the first 200,000 discarded as burn-in and the remainder thinned to one in every twenty iterations to save storage space. Model convergence was checked with R hat values⁴⁶. If the models did not converge with the initial conditions, we increased iterations up to 5,000,000 (with the first 1,000,000 discarded and the remainder thinned to one in every 800). We decided to remove grid cells in which parameter estimates did not converge even with the increased iterations, although the number of removed cells was very small (median of 2.5 grid cells in 20 out of the 775 (2.6%) regional populations).

To estimate the population-level change in abundance since 1990 for each species in a particular grid cell, we first regressed the estimates of the cell-specific year effect $\beta_{j(i),t}$ in every posterior sample against years. To account for uncertainty in slope estimates in this regression, we derived for every posterior sample a slope estimate from a normal distribution with the mean of the estimated mean slope and s.d. of the standard error of the slope. We then calculated the mean, median, variance and 2.5th and 97.5th percentiles of the estimated slopes from all posterior samples. We aggregated all estimates by species on the basis of definitions from BirdLife International³⁶. We used the mean and 2.5th and 97.5th percentiles of the estimated slopes for creating species-level maps (Fig. 1a–c and Supplementary Data 1). To calculate community-level changes in abundance (Fig. 2a) and community-level changes for species with different migratory statuses (Extended Data Fig. 2), we used the mean slopes across all species or all species in a particular group observed in each grid cell, weighted by the inverse of slope variance in each species to account for uncertainties. To further calculate mean community-level changes in each region (Fig. 2b), we used the mean of the community-level changes across all grid cells in each region, weighted by the inverse of associated variance.

Driver analysis. We first tested correlations among the nine spatial explanatory variables in 2,079 $1^\circ \times 1^\circ$ grid cells that had abundance change estimates (Extended Data Table 2). GDP per capita and governance were relatively strongly correlated ($r = 0.76$) with one another. Thus, considering that GDP growth rates are another measure of economic growth, we decided to exclude GDP per capita from the main analyses; instead, we tested its effect in a separate set of analyses in which governance was replaced with GDP per capita. In these analyses, considering the hypothesized nonlinear relationship between GDP per capita and species abundance changes (Extended Data Table 1), we used linear and quadratic terms of GDP per capita. We present the results of these analyses that use GDP per capita in Extended Data Fig. 5 and Supplementary Discussion.

To identify factors associated with waterbird abundance changes at the community, species and population levels, we conducted two types of analyses, both of which were implemented with WinBUGS v.1.4.3 and the R2WinBUGS package in R v.3.3.2.

In the first analysis, in which the response variable was community-level changes in abundance within each grid cell (Fig. 2a), we used a CAR model:

$$\mu_i = \alpha + \beta \mathbf{X}_i + \gamma_i$$

where the community-level change r_i in cell i was assumed to derive from a normal distribution with mean μ_i and variance σ_μ^2 ; β represents the vector of regression coefficients and \mathbf{X}_i the vector of explanatory variables. On the basis of the hypotheses shown in Extended Data Table 1, we used eight explanatory variables in each grid cell: surface water change, GDP growth rates, changes in human population density, crop area, temperature, and precipitation, protected area coverage and governance. We tested interaction terms between latitude and temperature change, and latitude and precipitation change, as population responses to temperature and precipitation can vary by latitude⁴⁷. We also tested a third interaction term between governance and protected area coverage, because governance can affect the effectiveness of conservation efforts⁴⁸. All explanatory variables were standardized before model fitting. The spatially-correlated random effect γ_i used an intrinsic Gaussian CAR prior distribution with variance σ_γ^2 , as described in equation (2). Prior distributions of parameters were set as non-informatively as possible; we used gamma distributions with a mean of 1 and variance of 1,000 for the inverse of σ_μ^2 and σ_γ^2 , normal distributions with a mean of 0 and variance of 1,000 for β_j , and an improper uniform distribution (a uniform distribution on an infinite interval) for the intercept α , as recommended by a previous study⁴⁹. Each MCMC algorithm was run with three chains with different initial values for 1,000,000 iterations, with the first 500,000 discarded as burn-in and the remainder thinned to one in every 100 iterations to save storage space. Model convergence was checked with R hat values.

Next, for 293 species observed in ten or more grid cells, we adopted the within-subject centring approach⁵⁰ under a hierarchical modelling framework to explicitly distinguish species-level effects (explaining variations in species-level abundance changes between species) and population-level effects (explaining variations in population-level abundance changes within species) of explanatory variables. In this model, the species effect μ_s , representing the species-level change

in abundance of species s , is drawn from a normal distribution with a mean of ν_s and variance of $\sigma_{\nu_s}^2$. The variable ν_s is further modelled with species-level explanatory variables:

$$\nu_s = \alpha + \sum_{k=1}^9 \beta_{Bk} \bar{x}_{k,s} + \sum_{k=10}^{12} \beta_{Bk} z_{k,s} + \eta_s$$

where α is the global intercept and β_{Bk} represents the species-level effect. The mean of spatial explanatory variable k across all grid cells where species s was recorded is represented by $\bar{x}_{k,s}$. Even if the estimated species-level abundance changes are biased owing to geographical biases in available grid cells, they match up with $\bar{x}_{k,s}$ because the calculation of both variables is performed on the same set of grid cells. The spatial explanatory variables used were derived from the hypotheses in Extended Data Table 1; we dropped changes in human population density and crop area, as these were least influential in the analysis of community-level population changes and also in a preliminary analysis of this species-level model. We therefore used the remaining six explanatory variables (surface water change, GDP growth rates, changes in temperature, changes in precipitation, protected area coverage and governance) and the same three interaction terms as used in the community-level analysis. The term $z_{k,s}$ represents three explanatory variables in species characteristics, described in Extended Data Table 1. The random term η_s accounts for phylogenetic dependence among species and is drawn from a multivariate normal distribution (MVN)^{51,52}:

$$\begin{aligned} \eta_s &\sim \text{MVN}(0, \delta^2 \Sigma) \\ \Sigma_{\lambda} &= \lambda \Sigma + (1 - \lambda) I \end{aligned} \quad (4)$$

where Σ is a scaled variance–covariance matrix calculated from an ultrametric phylogenetic tree. By scaling Σ to a height of one, we can interpret δ^2 as the residual variance⁵¹. To enable the strength of phylogenetic signal to vary, we also incorporated Pagel's λ ^{53,54} into the matrix in equation (4) with the identity matrix I . Here λ is a coefficient that multiplies the off-diagonal elements of Σ ; a λ close to zero implies that the phylogenetic signal in the data is low, which suggests independence in the error structure of the data points, whereas a λ that is close to one suggests a good agreement with the Brownian motion evolution model and thus suggests correlation in the error structure^{51,54}. To incorporate uncertainties⁵⁵ in phylogenetic trees in the calculation of Σ , we used a sample of 100 trees from a comprehensive avian phylogeny⁵⁶ as the prior distribution for our analysis⁵¹. More specifically, one of the 100 trees was randomly drawn in each iteration and used for the calculation of Σ .

The population-level change in abundance $r_{s,i}$ of species s in grid cell i was then assumed to derive from a normal distribution with mean $\mu_{s,i}$ and variance $\sigma_{\mu_s}^2$, where $\mu_{s,i}$ is modelled using the species effect μ_s :

$$\mu_{s,i} = \mu_s + \sum_{j=1}^6 \beta_{Wsj} (x_{j,i} - \bar{x}_{j,s}) + \gamma_{s,i}$$

Here β_{Wsj} represents the population-level effect for species s , explaining within-species variations in population-level abundance changes ($\mu_{s,i} - \mu_s$) by within-species variations in explanatory variables ($x_{j,i} - \bar{x}_{j,s}$); here, $x_{j,i}$ is the explanatory variable j in grid cell i and $\bar{x}_{j,s}$ is the mean of x_j for species s . The species-specific β_{Wsj} is the random effect governed by hyper-parameters as:

$$\beta_{Wsj} \sim \text{normal}(h\beta_{Wj}, \sigma_{\beta_{Wj}}^2)$$

For population-level effects, we used the six explanatory variables (surface water change, GDP growth rates, changes in temperature, changes in precipitation, protected area coverage and governance). Spatial autocorrelation within each species is accounted for by $\gamma_{s,i}$, which is drawn from an intrinsic Gaussian CAR prior distribution with variance $\sigma_{\gamma_s}^2$ as in equation (2).

As non-informative prior distributions, we used a gamma distribution with a mean of 1 and variance of 100 for $\sigma_{\nu_s}^2$, δ^2 , $\sigma_{\mu_s}^2$, $\sigma_{\beta_{Wj}}^2$ and $\sigma_{\gamma_s}^2$, uniform distribution on the interval [0, 1] for λ , normal distributions with a mean of 0 and variance of 100 for α , β_{Bk} , and $h\beta_{Wj}$. Each MCMC algorithm was run with three chains with different initial values for 10,000 iterations with the first 5,000 discarded as burn-in and the remainder thinned to one in every two iterations to save storage space. Model convergence was checked with R hat values. Owing to differences in the definition of species between the two sources used^{36,56}, in four cases we combined two separate species defined in the BirdLife Checklist³⁶ into one for the species-level analysis. These were the Kentish plover *Charadrius alexandrinus* and snowy plover *C. nivosus*, common snipe *Gallinago gallinago* and Wilson's snipe *G. delicata*, European herring gull *Larus argentatus* and Arctic herring gull

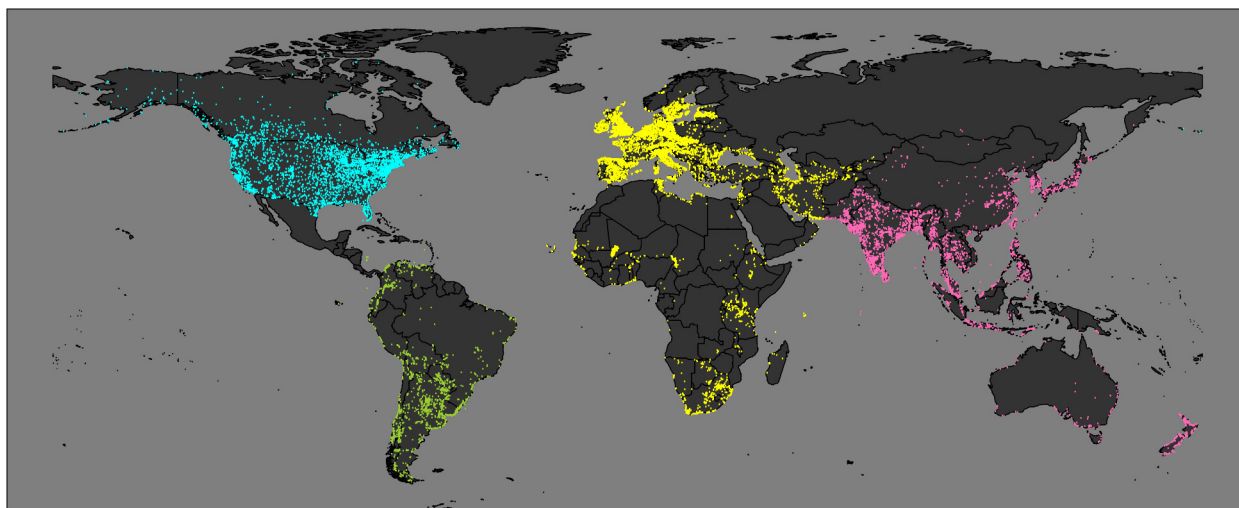
L. smithsonianus, and common moorhen *Gallinula chloropus* and common gallinule *G. galeata*.

Code availability. All the R and WinBUGS codes used for the analyses are available from the corresponding author upon request.

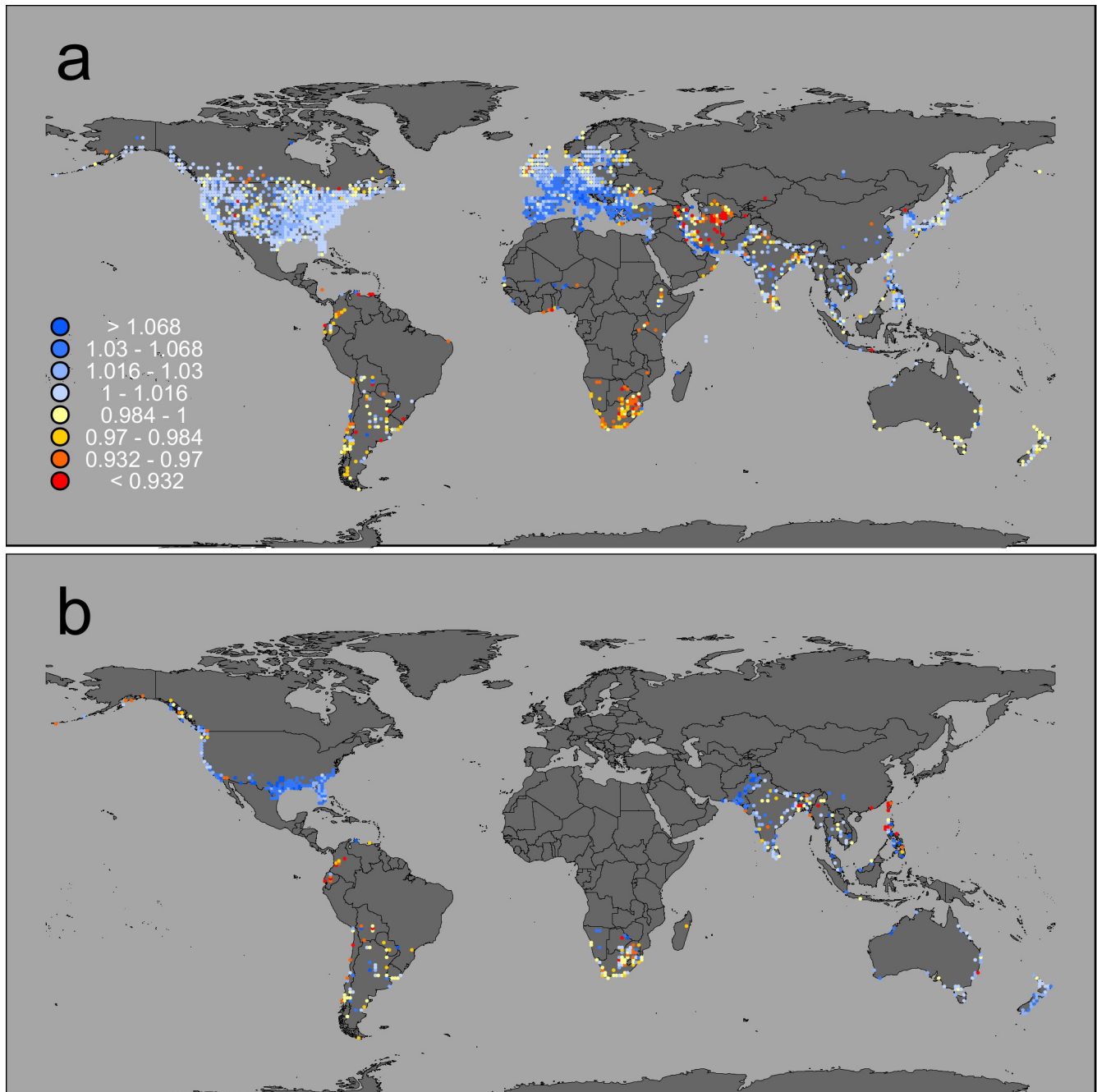
Data availability. The waterbird count data used in this study are collated and managed by Wetlands International and the National Audubon Society, and are available on request. All maps in figures are derived from the Natural Earth dataset (v.1.4.0) at 1:110 m scale (<http://www.naturalearthdata.com/downloads/110m-cultural-vectors/110m-admin-0-countries/>). All the data that pertain to explanatory variables are freely available, as specified in Extended Data Table 1. Supplementary Data 1 is available at <https://doi.org/10.6084/m9.figshare.5669827>. Supplementary Data 2 is available in the online version of the paper.

29. Delany, S. *Guidance on Waterbird Monitoring Methodology: Field Protocol for Waterbird Counting* (Wetlands International, 2010).
30. Dunn, E. H. et al. Enhancing the scientific value of the Christmas Bird Count. *Auk* **122**, 338–346 (2005).
31. van Roomen, M. et al. *Waterbird and Site Monitoring along the Atlantic Coast of Africa: Strategy and Manual* (BirdLife International, Common Wadden Sea Secretariat and Wetlands International, 2014).
32. LeBaron, G. S. *The 115th Christmas Bird Count* (National Audubon Society, 2015).
33. van Roomen, M., van Winden, E. & van Turnhout, C. *Analyzing Population Trends at the Flyway Level for Bird Populations Covered by the African Eurasian Waterbird Agreement: Details of a Methodology* (SOVON Dutch Centre for Field Ornithology, 2011).
34. Kaufmann, D., Kraay, A. & Mastruzzi, M. *The Worldwide Governance Indicators: Methodology and Analytical Issues* (September 2010) <https://ssrn.com/abstract=1682130> (2010).
35. Hsu, A. et al. 2016 *Environmental Performance Index* <http://epi.yale.edu/reports/2016-report> (2016).
36. BirdLife International. *The BirdLife Checklist of the Birds of the World: Version 7* http://www.birdlife.org/datazone/userfiles/file/Species/Taxonomy/BirdLife_Checklist_Version_70.zip (2014).
37. BirdLife International and NatureServe. *Bird Species Distribution Maps of the World* (BirdLife International and NatureServe, 2014).
38. Gill, F. & Donsker, D. (eds) *IOC World Bird List (v 5.1)* http://www.worldbirdnames.org/DOI-5/master_ioc_list_v5.1.xls (2015).
39. Amano, T., Okamura, H., Carrizo, S. F. & Sutherland, W. J. Hierarchical models for smoothed population indices: the importance of considering variations in trends of count data among sites. *Ecol. Indic.* **13**, 243–252 (2012).
40. Amano, T., Székely, T., Koyama, K., Amano, H. & Sutherland, W. J. A framework for monitoring the status of populations: an example from wader populations in the East Asian–Australasian flyway. *Biol. Conserv.* **143**, 2238–2247 (2010).
41. Link, W. A. & Sauer, J. R. Seasonal components of avian population change: joint analysis of two large-scale monitoring programs. *Ecology* **88**, 49–55 (2007).
42. Lunn, D. J., Thomas, A., Best, N. & Spiegelhalter, D. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat. Comput.* **10**, 325–337 (2000).
43. Sturtz, S., Ligges, U. & Gelman, A. R2WinBUGS: a package for running WinBUGS from R. *J. Stat. Softw.* **12**, 1–16 (2005).
44. R Core Team. *R: a Language and Environment for Statistical Computing* <http://www.R-project.org/> (R Foundation for Statistical Computing, 2016).
45. Link, W. A., Sauer, J. R. & Niven, D. K. A hierarchical model for regional analysis of population change using Christmas Bird Count data, with application to the American Black Duck. *Condor* **108**, 13–24 (2006).
46. Gelman, A., Carlin, J., Stern, H. & Rubin, D. *Bayesian Data Analysis* 2nd edn (Chapman & Hall and CRC, 2003).
47. Pearce-Higgins, J. W. et al. Geographical variation in species' population responses to changes in temperature and precipitation. *Proc. R. Soc. Lond. Ser. B* **282**, 20151561 (2015).
48. Bare, M., Kauffman, C. & Miller, D. C. Assessing the impact of international conservation aid on deforestation in sub-Saharan Africa. *Environ. Res. Lett.* **10**, 125010 (2015).
49. Thomas, A., Best, N., Lunn, D., Arnold, R. & Spiegelhalter, D. *GeoBUGS User Manual* <http://www.mrc-bsu.cam.ac.uk/software/bugs/> (2004).
50. van de Pol, M. & Wright, J. A simple method for distinguishing within- versus between-subject effects using mixed models. *Anim. Behav.* **77**, 753–758 (2009).
51. de Villemereuil, P., Wells, J. A., Edwards, R. D. & Blomberg, S. P. Bayesian models for comparative analysis integrating phylogenetic uncertainty. *BMC Evol. Biol.* **12**, 102 (2012).
52. Abadi, F. et al. Importance of accounting for phylogenetic dependence in multi-species mark-recapture studies. *Ecol. Modell.* **273**, 236–241 (2014).
53. Pagel, M. Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884 (1999).
54. Freckleton, R. P., Harvey, P. H. & Pagel, M. Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.* **160**, 712–726 (2002).
55. Donoghue, M. J. & Ackerly, D. D. Phylogenetic uncertainties and sensitivity analyses in comparative biology. *Philos. Trans. R. Soc. Lond. B* **351**, 1241–1249 (1996).
56. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).
57. Grossman, G. M. & Krueger, A. B. Economic growth and the environment. *Q. J. Econ.* **110**, 353–377 (1995).

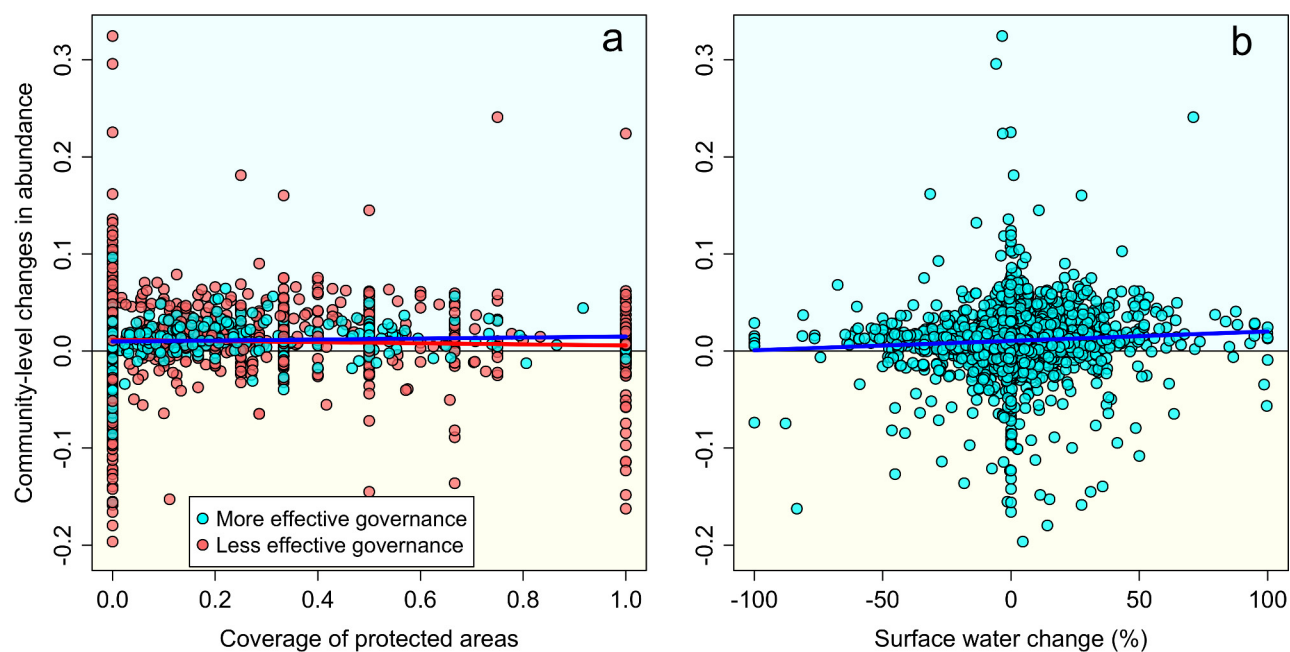
58. Cardillo, M. *et al.* Human population density and extinction risk in the world's carnivores. *PLoS Biol.* **2**, e197 (2004).
59. McKee, J., Chambers, E. & Guseman, J. Human population density and growth validated as extinction threats to mammal and bird species. *Hum. Ecol.* **41**, 773–778 (2013).
60. Center for International Earth Science Information Network - CIESIN - Columbia University, and Centro Internacional de Agricultura Tropical - CIAT. *Gridded Population of the World, Version 3 (GPWv3): Population Density Grid* <http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-density> (2005).
61. Green, R. E., Cornell, S. J., Scharlemann, J. P. W. & Balmford, A. Farming and the fate of wild nature. *Science* **307**, 550–555 (2005).
62. Friedl, M. A. *et al.* MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **114**, 168–182 (2010).
63. Stephens, P. A. *et al.* Consistent response of bird populations to climate change on two continents. *Science* **352**, 84–87 (2016).
64. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 dataset. *Int. J. Climatol.* **34**, 623–642 (2014).
65. Kleijn, D., Cherkaoui, I., Goedhart, P. W., van der Hout, J. & Lammertsma, D. Waterbirds increase more rapidly in Ramsar-designated wetlands than in unprotected wetlands. *J. Appl. Ecol.* **51**, 289–298 (2014).
66. Pavón-Jordán, D. *et al.* Climate-driven changes in winter abundance of a migratory waterbird in relation to EU protected areas. *Divers. Distrib.* **21**, 571–582 (2015).
67. UNEP-WCMC and IUCN. *Protected Planet: The World Database on Protected Areas (WDPA)* www.protectedplanet.net (2015).
68. Mace, G. M. *et al.* Quantification of extinction risk: IUCN's system for classifying threatened species. *Conserv. Biol.* **22**, 1424–1442 (2008).
69. Sanderson, F. J., Donald, P. F., Pain, D. J., Burfield, I. J. & van Bommel, F. P. J. Long-term population declines in Afro-Palearctic migrant birds. *Biol. Conserv.* **131**, 93–105 (2006).
70. Robbins, C. S., Sauer, J. R., Greenberg, R. S. & Droege, S. Population declines in North American birds that migrate to the neotropics. *Proc. Natl Acad. Sci. USA* **86**, 7658–7662 (1989).
71. Pocock, M. J. O. Can traits predict species' vulnerability? A test with farmland passerines in two continents. *Proc. R. Soc. Lond. Ser. B* **278**, 1532–1538 (2011).
72. Owens, I. P. F. & Bennett, P. M. Ecological basis of extinction risk in birds: habitat loss versus human persecution and introduced predators. *Proc. Natl Acad. Sci. USA* **97**, 12144–12148 (2000).
73. Wilman, H. *et al.* EltonTraits 1.0: species-level foraging attributes of the world's birds and mammals. *Ecology* **95**, 2027 (2014).



Extended Data Figure 1 | Distribution of the 25,769 survey sites used in the analyses. Sites from the International Waterbird Census are shown in yellow (African–Eurasian Waterbird Census), pink (Asian Waterbird Census) and green (Neotropical Waterbird Census). Christmas Bird Count shown in cyan.



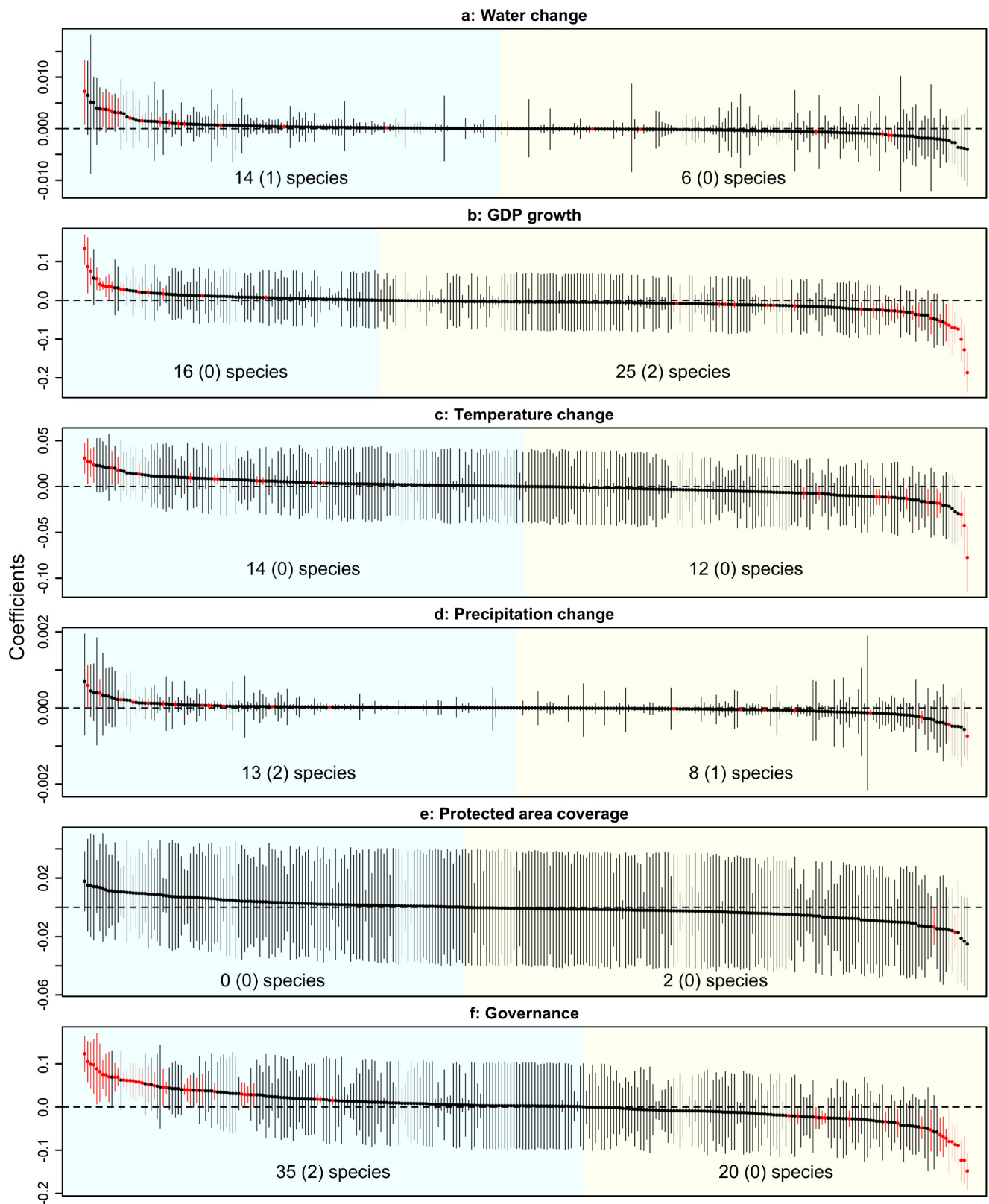
Extended Data Figure 2 | Global distribution of mean annual changes in abundance. a, b, Mean annual changes in abundance for 373 migratory (a) and 88 non-migratory (b) waterbird species (that is, community-level changes). The migratory status of each species was assigned using the BirdLife Data Zone (see Methods).



Extended Data Figure 3 | Relationships between community-level changes in abundance and protected areas or surface water.

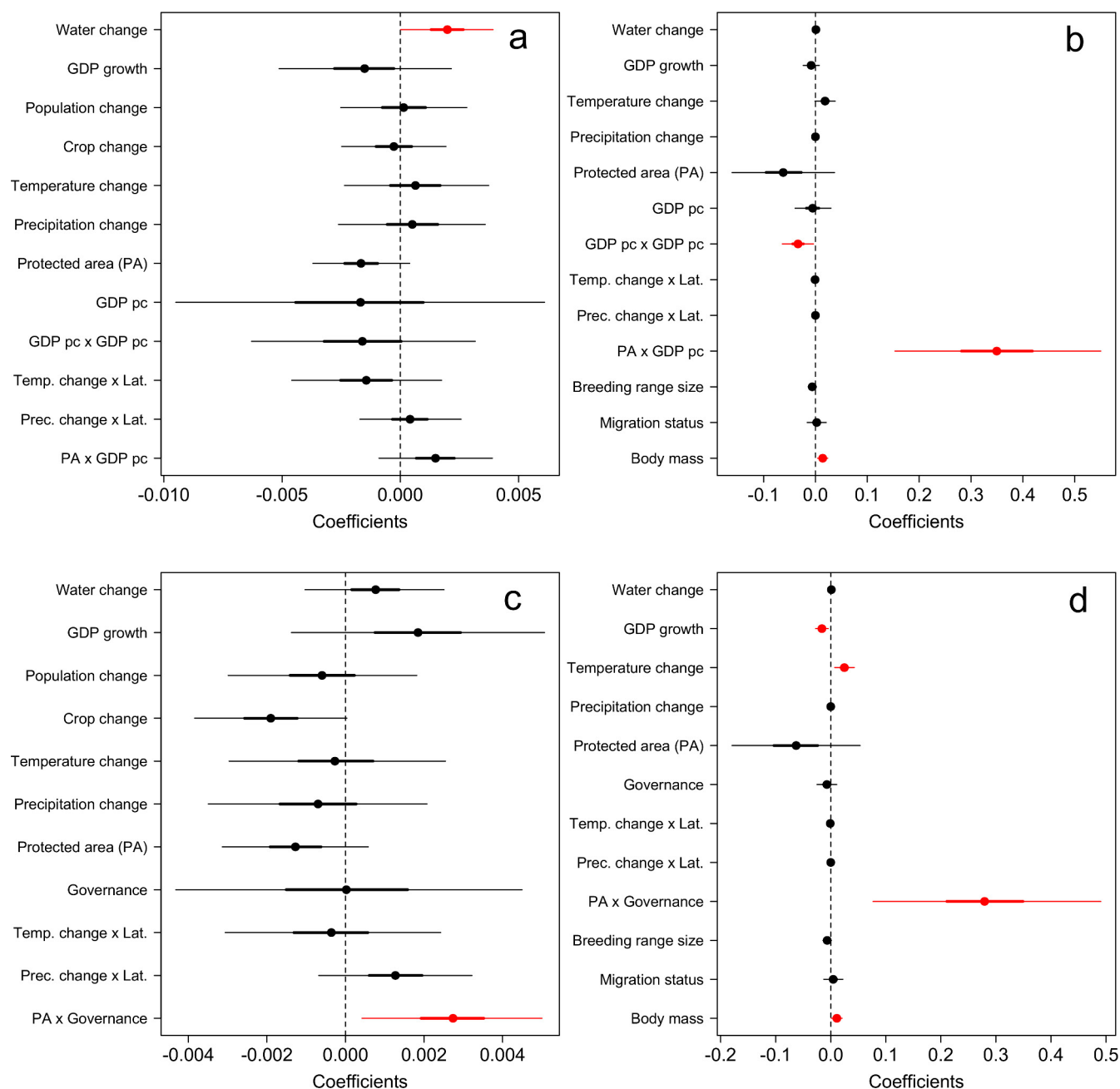
a, Relationship between community-level changes in abundance and the proportion of sites covered by protected areas. **b**, Relationship

between community-level changes in abundance and surface water change. Regression lines are based on the estimated coefficients in Fig. 3a; values and regression lines for grid cells in areas with more (in blue) and less (in red) effective governance in **a**. $n = 2,079$ grid cells.



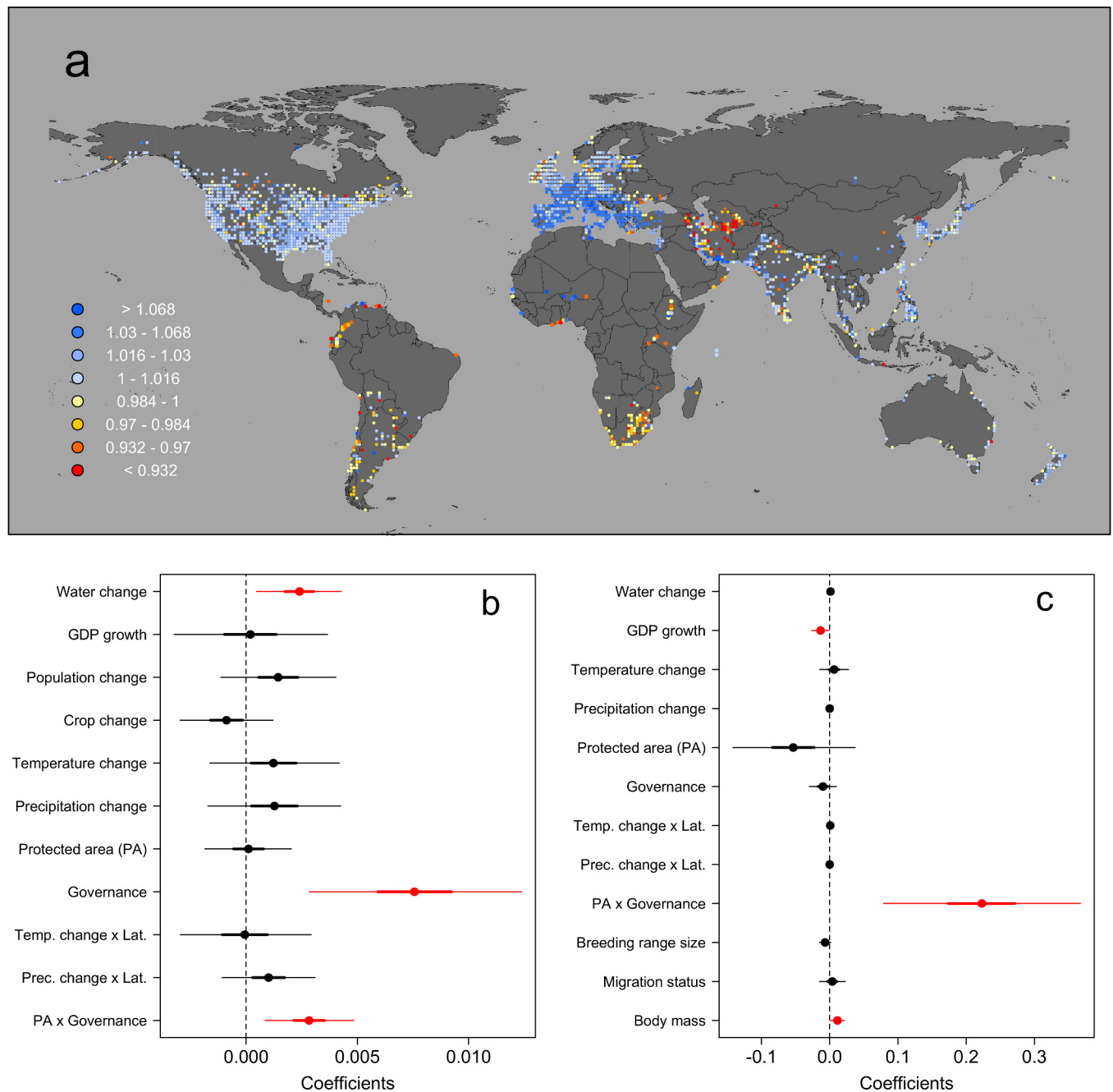
Extended Data Figure 4 | Effects of six hypothesized predictors on population-level changes in abundance. a–f, Medians and 95% credible intervals of the estimated coefficients for 293 species are shown in order of decreasing positive effect size from the left (those with 95% credible

intervals not overlapping with zero shown in red). The numbers of species with significant positive and negative coefficients are also shown, with the number of non-migratory species in parentheses. See Extended Data Table 1 for more detail regarding predictors.



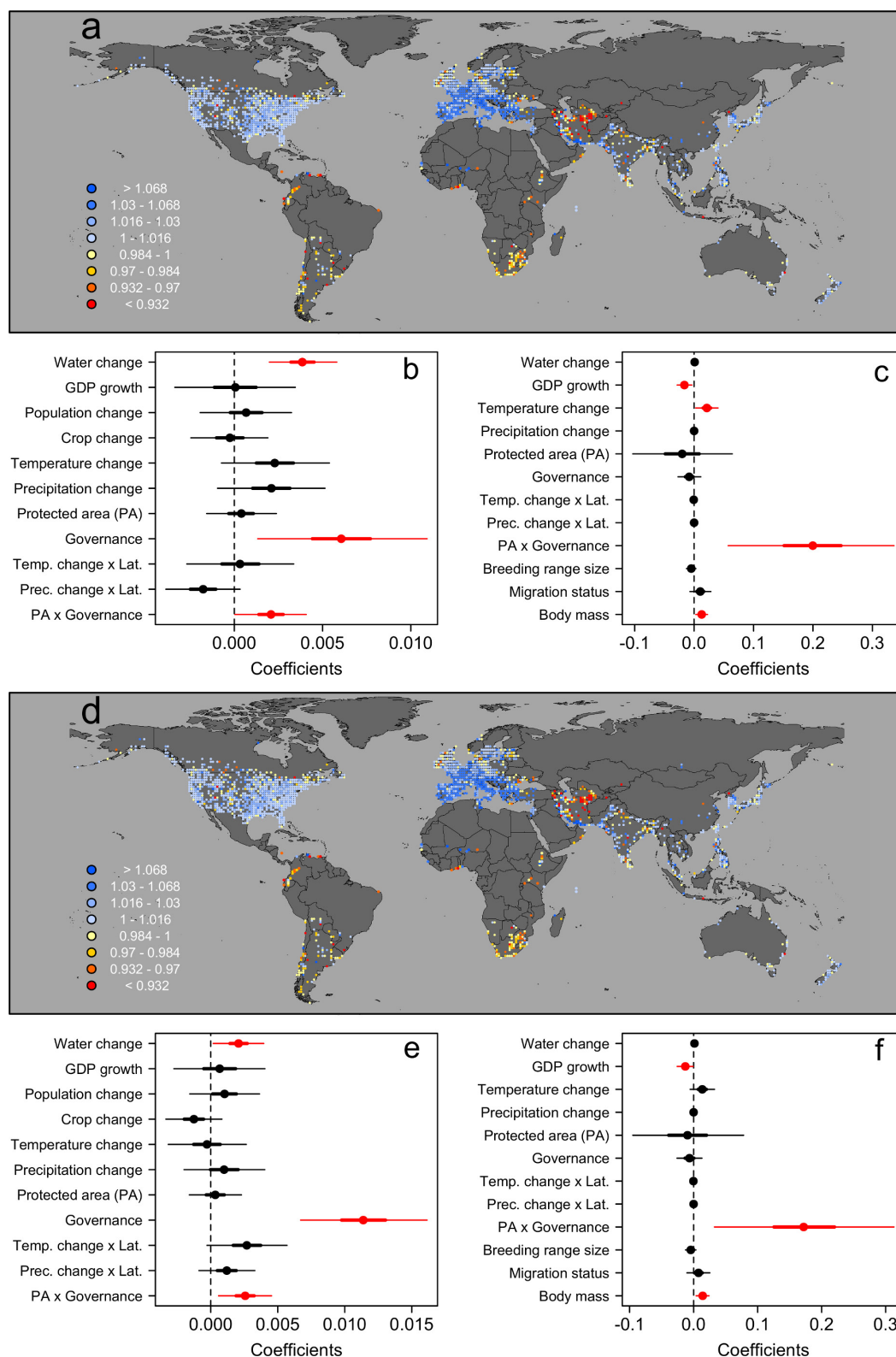
Extended Data Figure 5 | Sensitivity of results to the correlation between governance and GDP per capita and designation years of protected areas. **a, b**, Estimated coefficients in the multivariate analysis of community-level ($n = 2,079$ grid cells) (**a**) and species-level (on the basis of 293 species; see Supplementary Data 2 for the number of grid cells for each species) (**b**) changes in abundance, in which governance was replaced by linear and quadratic terms of GDP per capita. **c, d**, Estimated coefficients in the multivariate analysis of

community-level ($n = 2,079$ grid cells) (**c**) and species-level (on the basis of 293 species; see Supplementary Data 2 for the number of grid cells in each species) (**d**) changes in abundance, in which only protected areas known to have been designated before 1990 (the first survey year in our dataset) were used (most conservative approach). Posterior medians with 95% and 50% (thick lines) credible intervals are shown. Coefficients with 95% credible intervals not overlapping with zero are shown in red.



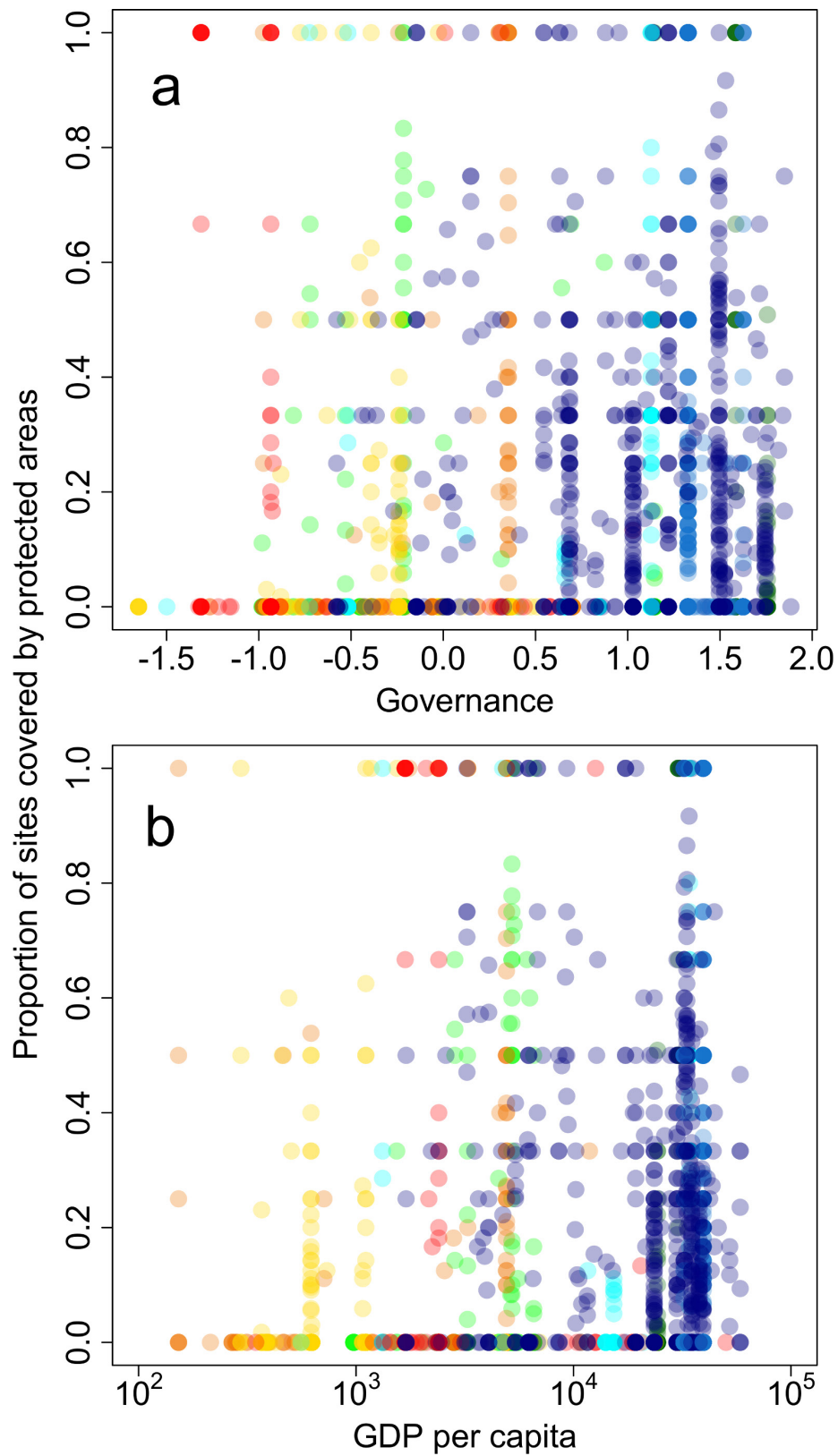
Extended Data Figure 6 | Sensitivity of the results to the inclusion of seabird species. **a**, Global distribution of mean annual changes in abundance across 447 waterbird species, excluding the 14 seabird species, between 1990 and 2013. **b**, **c**, Estimated coefficients in the multivariate analysis of community-level ($n = 2,079$ grid cells) (**b**) and species-level

(on the basis of 447 species; see Supplementary Data 2 for the number of grid cells in each species) (**c**) changes in abundance, in which the 14 seabird species were excluded. Posterior medians with 95% and 50% (thick lines) credible intervals are shown. Coefficients with 95% credible intervals not overlapping with zero are shown in red.



Extended Data Figure 7 | Sensitivity of the results to the choice of CBC survey sites for the analyses. **a**, Global distribution of mean annual changes in abundance across 461 waterbird species between 1990 and 2013, after excluding 41 CBC grid cells that contained neither landscape-scale wetland areas nor local-scale surface water occurrences within 1km of all the survey sites included. **b**, **c**, Estimated coefficients in the multivariate analysis of community-level ($n = 2,038$ grid cells) (**b**) and species-level (on the basis of 293 species) (**c**) changes in abundance, in which 41 CBC grid cells that contained neither landscape-scale wetland areas nor local-scale surface water occurrences within 1km of all the

survey sites were excluded. **d**, Global distribution of mean annual changes in abundance across 461 waterbird species between 1990 and 2013, after excluding eight CBC grid cells in which the proportion of urban areas was over 0.3. **e**, **f**, Estimated coefficients in the multivariate analysis of community-level ($n = 2,071$ grid cells) (**e**) and species-level (on the basis of 293 species) (**f**) changes in abundance, in which eight CBC grid cells with a proportion of urban areas of over 0.3 were excluded. Posterior medians with 95% and 50% (thick lines) credible intervals are shown. Coefficients with 95% credible intervals not overlapping with zero are shown in red.



Extended Data Figure 8 | Relationships between the proportion of sites covered by protected areas and governance or GDP per capita.
a, b, The relationship between governance (a) or GDP per capita (b) and the proportion of sites covered by protected areas. Colours indicate

regions: blue, North America; green, South America; navy, Europe; orange, Africa; red, western and central Asia; yellow, south and southeast Asia; cyan, east Asia and Russia; and dark green, Oceania.

Extended Data Table 1 | Hypotheses and explanatory variables tested for explaining the patterns in waterbird abundance changes over space and species

Hypotheses	Drivers	Descriptions	Explanatory variables used	Data sources
<i>Anthropogenic impacts</i>	Surface water	Surface water provides an essential habitat for most wetland-dependent species ¹ , thus its decline can threaten the status of waterbirds	Mean changes (%) in surface water occurrence between 1984-1999 and 2000-2015, within 1km from each survey site	Global Surface Water ¹⁶
	Economic growth	Economic growth poses a threat to species through habitat loss and degradation but can also improve environmental quality at a high economic level ⁵⁷ .	Mean country-level GDP per capita between 1990 and 2010	World Bank*
			Mean country-level GDP growth rate (annual %) between 1990 and 2010	World Bank†
	Human population growth	High species extinction risk is associated with high human population density ⁵⁸ and rapid human population growth ⁵⁹ .	Mean changes in human population density between 1990 and 2000	Population Density Grid v3 ⁶⁰
	Agricultural expansion	Farming is the biggest source of threats to bird species ⁶¹ .	Changes in crop area (croplands and cropland/natural vegetation mosaics) between 2001 and 2010	Collection 5 MODIS Global Land Cover Type product ⁶²
<i>Conservation efforts and effectiveness</i>	Climate change	Climate change is a strong predictor of bird abundance changes ⁶³ .	Changes in mean Dec-Feb temperature between 1985-1990 and 2005-2010	CRU TS3.10 Dataset ⁶⁴
			Changes in mean Dec-Feb precipitation between 1985-1990 and 2005-2010	CRU TS3.10 Dataset ⁶⁴
	Protected areas	Waterbird abundance increased more rapidly in protected than in unprotected wetlands ^{65,66} .	Proportion of sites covered by protected areas	World Database on Protected Areas ⁶⁷
	Governance	Ineffective governance in a country is associated with species population declines ¹⁷ .	Mean of six country-level Worldwide Governance Indicators between 1996 and 2010	World Bank‡
<i>Species characteristics</i>	Geographical range size	Species with small geographical range may be more susceptible to large-scale, stochastic threats ⁶⁸ .	Breeding/resident geographical range size (km ²)	Birdlife Data Zone§
	Migratory status	Migratory species can be affected by conditions at multiple locations, thus tend to show population declines ^{69,70} .	Migrant or non-migrant	Birdlife Data Zone§
	Body size	Body size is a strong predictor of bird abundance changes ⁷¹ but its association with bird extinction risk can be both positive and negative, depending on threats to the species ⁷²	Body mass (g)	EltonTraits 1.0 ⁷³

*<http://data.worldbank.org/indicator/NY.GDP.PCAP.KD>.†<http://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG>.‡<http://data.worldbank.org/data-catalog/worldwide-governance-indicators>.§<http://datazone.birdlife.org/home>.

Extended Data Table 2 | Correlation matrix (Spearman's rank correlation) of nine potential predictors of waterbird abundance changes ($n=2,079$ grid cells)

	GDP per capita	Water change	GDP growth rate	Human population change	Crop area change	Dec-Feb temperature change	Dec-Feb precipitation change	Protected area coverage
Water change	-0.087							
GDP growth rate	-0.502	0.003						
Human population change	-0.326	-0.047	0.442					
Crop area change	-0.095	0.039	0.208	0.140				
Dec-Feb temperature change	-0.176	-0.070	0.158	0.100	-0.087			
Dec-Feb precipitation change	0.005	0.045	-0.059	-0.043	-0.091	0.031		
Protected area coverage	0.002	0.002	-0.225	-0.077	-0.051	-0.121	-0.081	
Governance	0.755	-0.100	-0.547	-0.344	-0.169	-0.200	-0.086	0.047

Gross domestic product (GDP) per capita is given as \log_{10} -transformed values. Strong correlations ($|r| > 0.7$) are shown in bold.

Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans

J. Víctor Moreno-Mayar^{1*}, Ben A. Potter^{2*}, Lasse Vinner^{1*}, Matthias Steinrücken^{3,4}, Simon Rasmussen⁵, Jonathan Terhorst^{6,7}, John A. Kamm^{6,8}, Anders Albrechtsen⁹, Anna-Sapfo Malaspinas^{1,10,11}, Martin Sikora¹, Joshua D. Reuther², Joel D. Irish¹², Ripan S. Malhi^{13,14}, Ludovic Orlando¹, Yun S. Song^{6,15,16}, Rasmus Nielsen^{1,6,17}, David J. Meltzer^{1,18} & Eske Willerslev^{1,8,19}

Despite broad agreement that the Americas were initially populated via Beringia, the land bridge that connected far northeast Asia with northwestern North America during the Pleistocene epoch, when and how the peopling of the Americas occurred remains unresolved^{1–5}. Analyses of human remains from Late Pleistocene Alaska are important to resolving the timing and dispersal of these populations. The remains of two infants were recovered at Upward Sun River (USR), and have been dated to around 11.5 thousand years ago (ka)⁶. Here, by sequencing the USR1 genome to an average coverage of approximately 17 times, we show that USR1 is most closely related to Native Americans, but falls basal to all previously sequenced contemporary and ancient Native Americans^{1,7,8}. As such, USR1 represents a distinct Ancient Beringian population. Using demographic modelling, we infer that the Ancient Beringian population and ancestors of other Native Americans descended from a single founding population that initially split from East Asians around 36 ± 1.5 ka, with gene flow persisting until around 25 ± 1.1 ka. Gene flow from ancient north Eurasians into all Native Americans took place 25–20 ka, with Ancient Beringians branching off around 22–18.1 ka. Our findings support a long-term genetic structure in ancestral Native Americans, consistent with the Beringian ‘standstill model’⁹. We show that the basal northern and southern Native American branches, to which all other Native Americans belong, diverged around 17.5–14.6 ka, and that this probably occurred south of the North American ice sheets. We also show that after 11.5 ka, some of the northern Native American populations received gene flow from a Siberian population most closely related to Koryaks, but not Palaeo-Eskimos¹, Inuits or Kets¹⁰, and that Native American gene flow into Inuits was through northern and not southern Native American groups¹. Our findings further suggest that the far-northern North American presence of northern Native Americans is from a back migration that replaced or absorbed the initial founding population of Ancient Beringians.

The details of the peopling of the Americas, and particularly the population history of Beringia, remain unresolved^{2,3}. Humans were present in the Americas south of the continental ice sheets by around 14.6 ka¹¹, indicating that they traversed Beringia earlier. During the Last Glacial Maximum (LGM), this region was marked by harsh climates and glacial barriers⁵, which may have led to the isolation of populations for extended periods, and at times complicated dispersal across the region¹². It remains unknown whether and for how long

Native American ancestors were isolated from Asian groups in Beringia before entering the Americas^{2,9,13}; whether one or more early migrations gave rise to the founding population of Native Americans^{1–4,7,14} (it is commonly agreed that the Palaeo-Eskimos and Inuit populations represent separate and later migrations^{1,15,16}); and when and where the basal split between southern and northern Native American (SNA and NNA, respectively) branches occurred. It also remains unresolved whether the genetic affinity between some SNA groups and indigenous Australasians^{2,3} reflects migration by non-Native Americans^{3,4,14}, early population structure within the first Americans³ or later gene flow². To resolve these uncertainties, a better understanding of the population history of Beringia, the entryway for the Pleistocene peopling of the Americas, is needed.

Genomic insight into that population history has now become available with the recently recovered infant remains (USR1 and USR2) from the Upward Sun River site, Alaska (eastern Beringia), which have been dated to approximately 11.5 ka^{6,17}. Mitochondrial DNA sequences (haplogroups C1 and B2, respectively) were previously acquired from these individuals^{6,17} (Supplementary Information sections 1, 4.5). We have since obtained whole-genome sequence data, which provide a broader opportunity to investigate the number, source(s) and structure of the initial founding population(s) and the timing and location of their subsequent divergence. We sequenced the genome of USR1 to an average depth of approximately 17×, on the basis of eight sequencing libraries from uracil-specific excision reagent-treated extracts that had previously been confirmed to contain DNA fragments with characteristic ancient DNA misincorporation patterns (Supplementary Information sections 2–4). We estimated modern human contamination to be around 0.14% based on the nuclear genome and about 0.15% based on mitochondrial DNA (Supplementary Information section 4). As expected, the error rate in the uracil-specific excision reagent-treated sequencing data was low (0.09% errors per base), and comparable to other high-coverage contemporary genomes, based on called genotypes (Supplementary Information section 4). Although USR2⁶ did not have sufficient endogenous DNA for high-coverage genome sequencing, we found that both individuals were close relatives (Supplementary Information section 5), equally related to worldwide present-day populations (Supplementary Fig. 4g).

We assessed the genetic relationship between USR1, a set of ancient genomes^{2,7,8,14,16} and a panel of 167 worldwide populations genotyped for 199,285 single-nucleotide polymorphisms^{1,2,18} (Supplementary

¹Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark. ²Department of Anthropology, University of Alaska, Fairbanks, Alaska 99775, USA. ³Department of Biostatistics and Epidemiology, University of Massachusetts, Amherst, Massachusetts 01003, USA. ⁴Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, USA. ⁵Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, 2800 Kongens Lyngby, Denmark. ⁶Department of Statistics, University of California, Berkeley, California 94720, USA. ⁷Department of Statistics, University of Michigan, Ann Arbor, Michigan 48109, USA. ⁸Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK. ⁹The Bioinformatics Centre, Department of Biology, University of Copenhagen, 2200 Copenhagen, Denmark. ¹⁰Department of Computational Biology, University of Lausanne, Lausanne, Switzerland. ¹¹Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. ¹²Research Centre in Evolutionary Anthropology and Palaeoecology, Liverpool John Moores University, Liverpool L3 3AF, UK. ¹³Department of Anthropology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. ¹⁴Carle R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. ¹⁵Computer Science Division, University of California, Berkeley, California 94720, USA. ¹⁶Chan Zuckerberg Biohub, San Francisco, California 94158, USA. ¹⁷Department of Integrative Biology, University of California, Berkeley, California 94720, USA. ¹⁸Department of Anthropology, Southern Methodist University, Dallas, Texas 75275, USA. ¹⁹Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK.

*These authors contributed equally to this work.

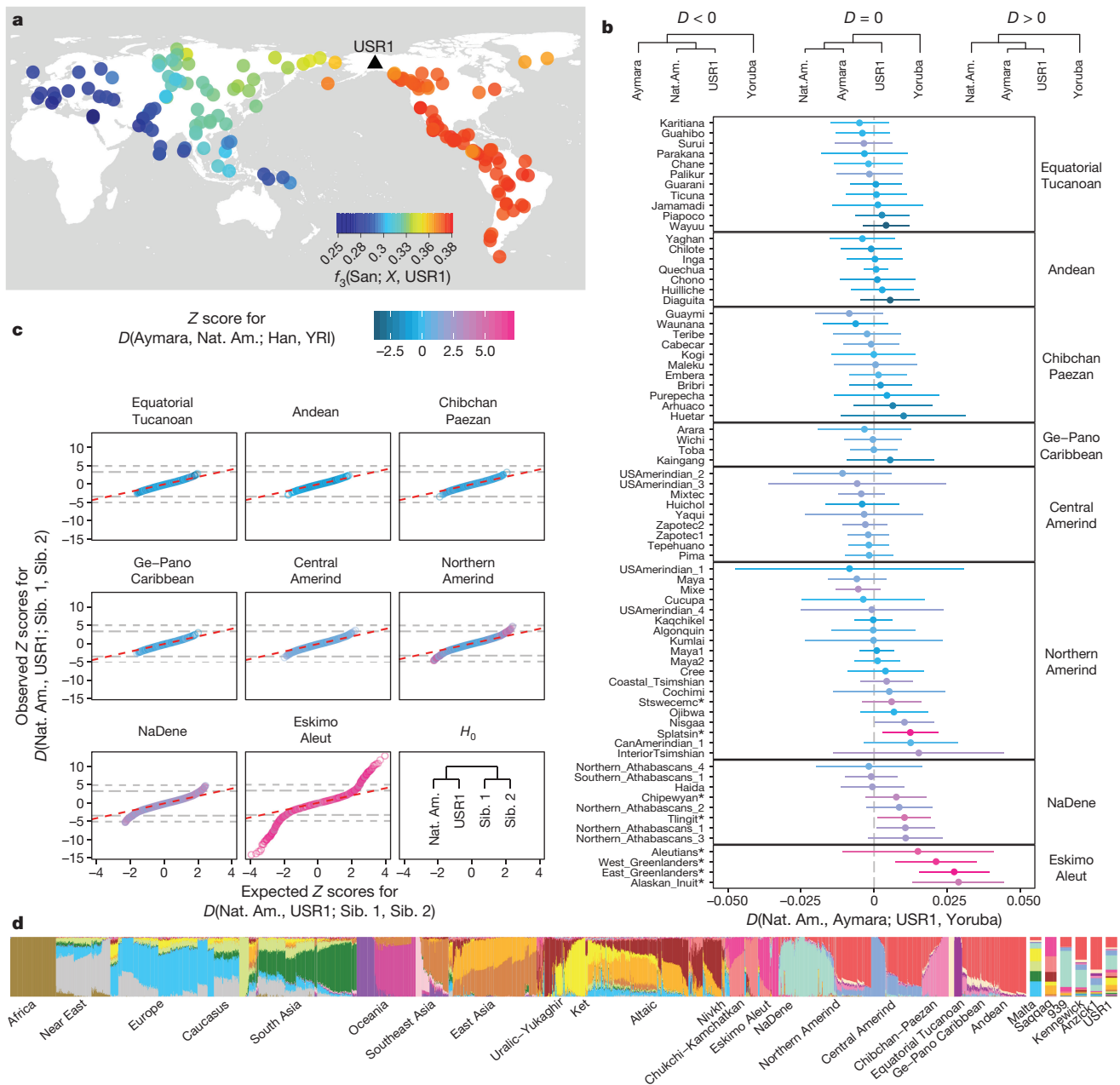


Figure 1 | Genetic affinities between USR1, present-day Native Americans and world-wide populations. a. f_3 statistics of the form $f_3(\text{San}; X, \text{USR1})$, for each population in the genotype panel. Warmer colours represent greater shared drift between a population (X) and USR1. **b.** D statistics of the form $D(\text{Native American, Aymara; USR1, Yoruba})$ (points). The Andean Aymara were used to represent SNA. *Native American populations with Asian admixture ($|Z|$ for $D(H1, \text{Aymara; Han, Yoruba}) > 3.3$) (Supplementary Fig. 5a). Error bars represent 1 and approximately 3.3 standard errors ($P \approx 0.001$). Native American populations were grouped by language family¹. **c.** Quantile-quantile plot comparing observed Z scores to the expected normal distribution under the null hypothesis (H_0), for all possible $D(\text{Native American, USR1; Siberian1, Siberian2})$. Colours correspond to the Z score obtained for $D(H1, \text{Aymara; Han, Yoruba})$. The expected normal distribution under the null hypothesis was computed for all groups jointly (Supplementary Information section 10.4). Thick and thin lines represent a Z score of approximately 3.3 ($P \approx 0.001$) and a Z score of approximately 4.91 ($P \approx 0.01$ after applying a Bonferroni correction for 11,322 tests). The bottom-right panel shows the expected tree under the null hypothesis. **d.** Admixture proportions estimated by ADMIXTURE²⁰ assuming $K = 20$ ancestral populations. Bars represent individuals, and colours represent admixture proportions from each ancestral component. Admixture proportions in ancient genomes (wider bars) were estimated using a genotype likelihood-based approach²¹. Nat. Am., Native American; Sib., Siberian.

Information section 6), using outgroup f_3 statistics¹⁹, model-based clustering^{20,21} and multidimensional scaling²² (Supplementary Information section 7–9). Outgroup f_3 statistics of the form $f_3(\text{San}; X, \text{USR1})$ revealed that USR1 is more closely related to present-day Native Americans than to any other tested population, followed by Siberian and East Asian populations^{1,2} (Fig. 1a). Pairwise comparisons of the f_3 statistics for USR1 and a set of ancient and contemporary Native American genomes^{2,7,14} (Supplementary Information section 6)

showed that all are similarly related to Eurasian, Australasian and African populations, although other Native American genomes (Aymara², Athabascan¹⁵, 939², Anzick¹⁷ and Kennewick¹⁴) have a higher affinity for contemporary Native Americans than does USR1 (Supplementary Information section 9). Multidimensional scaling and ADMIXTURE analysis showed that the USR1 genome did not cluster with any specific Native American group (Fig. 1d and Supplementary Fig. 3b). These results imply that USR1 belonged to a previously

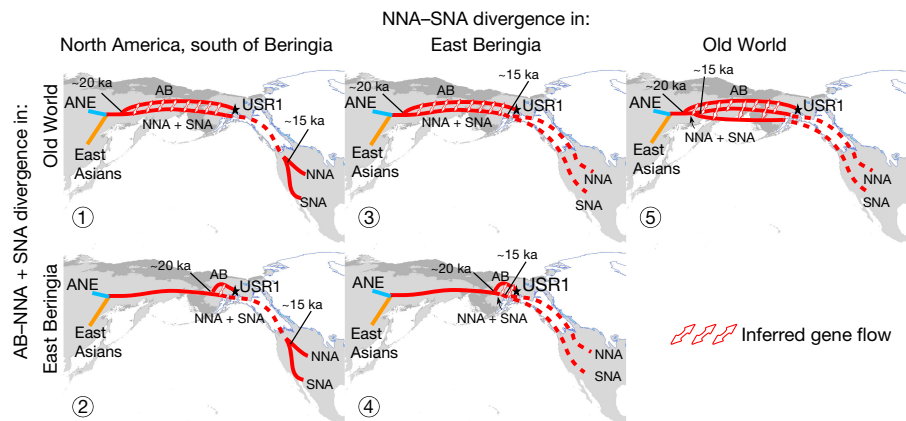


Figure 2 | Possible geographic locations for the USR1 and NNA-SNA splits. We propose two possible locations for the split between USR1 and other Native Americans: the Old World (scenarios 1, 3, 5) and Beringia (scenarios 2, 4); and three possible locations for the NNA-SNA split: the Old World (scenario 5), Beringia (scenarios 3, 4), and North America south of Beringia (scenarios 1, 2). Schematics show estimated glacial extent around 14.8 ka. Dashed lines represent the Native American

migration south of eastern Beringia, but they do not correspond to a specific migration route. Model discussion (Supplementary Information section 20) is based on extant archaeological evidence and inferred demographic parameters: a USR1–NNA and SNA split about 20 ka with ensuing moderate gene flow and a NNA–SNA split around 15 ka (Supplementary Information sections 18, 19). AB, Ancient Beringian; ANE, Ancient North Eurasian.

unknown Native American population that was not represented in the reference dataset, and which is herein identified as Ancient Beringians (Supplementary Information section 8.3).

To investigate whether USR1 derived from the same source population that gave rise to contemporary Native Americans, we computed 11,322 allele frequency-based D -statistics^{1,19} of the form $D(\text{Native American, USR1}; \text{Siberian1/Han, Siberian2/Han})$ (Supplementary Information section 10.4). The resulting Z -score distribution corresponds qualitatively to the expected normal distribution under the null hypothesis that USR1 forms a clade with Native Americans to the exclusion of Siberians and East Asians—except for a set of Eskimo–Aleut, Athabascan and Northern Amerind-speaking populations for which recent Asian gene flow has previously been documented^{1,2,14,18} (Fig. 1c and Supplementary Figs 5a, 6). Additionally, we found that present-day Native Americans and USR1 yield similar results for $D(\text{Native American/USR1, Han; Mal'ta, Yoruba})$, suggesting that they are equally related to the ancient north Eurasian population represented by the 24-thousand-year-old Mal'ta individual⁸ (Supplementary Information section 10.5). These results confirm that USR1 and present-day Native Americans derived from the same ancestral source, which carried a mixture of East Asian- and Mal'ta-related ancestry. We infer that descendants of this source represent the basal group that first migrated into the Americas.

To explore the relationship between USR1 and present-day Native Americans, we computed allele frequency-based and genome-wide D statistics of the form $D(\text{Native American, Aymara; USR1, Yoruba})$. We could not reject the null hypothesis that USR1 is an outgroup to any pair of Native Americans, with the exception of a set of populations bearing recent Asian gene flow^{1,2,14,18} (Fig. 1b and Supplementary Fig. 7). We confirmed the phylogenetic placement of USR1 at a basal position in the Native American clade using TreeMix²³ and two methods to estimate average genomic divergence and genetic drift, respectively (Supplementary Information sections 14–16). These results support the branching of USR1 within the Native American clade, but with USR1 being equidistant to NNA and SNA. Below we discuss the potential geographic locations of the split between USR1 and the common ancestor of NNA and SNA, and the NNA–SNA split (Fig. 2) on the basis of genetic results, the glacial geography of terminal Pleistocene North America^{24,25} and the extant archaeological evidence (Supplementary Information section 20).

Recent detection of an Australasian-derived genetic signature in some Native American groups^{2,3} led us to explore whether USR1 also bears this signature (Supplementary Information sections 10.7, 11–13).

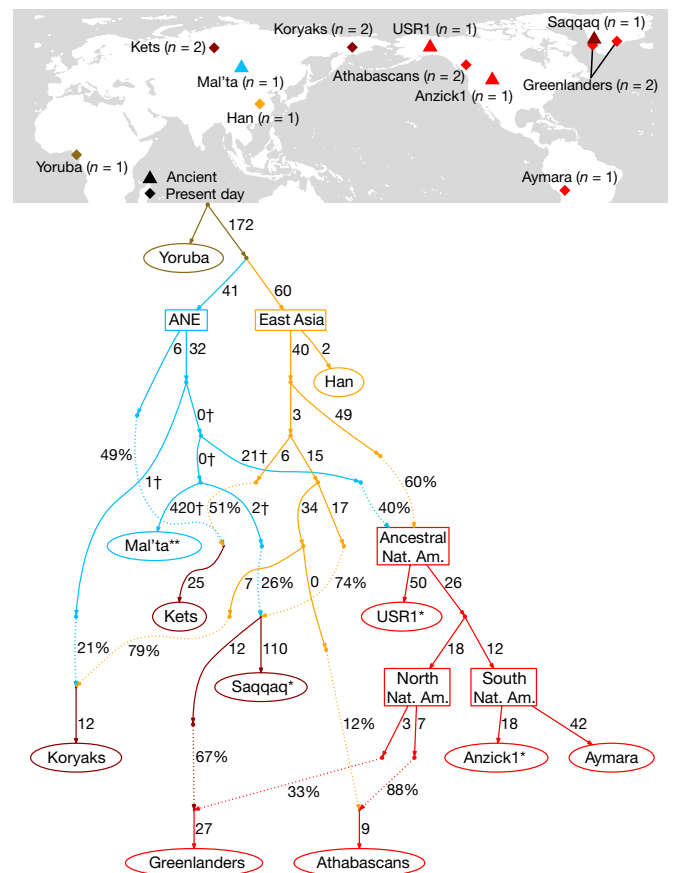


Figure 3 | A model for the formation of the different Native American populations. We fitted an admixture graph by sequentially adding admixed leaves to a 'seed' graph including the Yoruba, Han, Mal'ta, Ket, USR1, Anzick1 and Aymara genomes. For each 'non-seed' admixed group, we found the pair of edges that produced the best-fitting graph, based on the fitting and maximum $|Z|$ scores (3.27 for this graph). Ellipse-shaped nodes: sampled populations; box-shaped nodes: metapopulations. *Single high-depth ancient genome; **single low-depth genome. †Subgraphs with a structure that we were unable to resolve due to sequencing and genotyping error in the Saqqaq genome (Supplementary Information section 17). Sample sizes and locations are shown at the top.

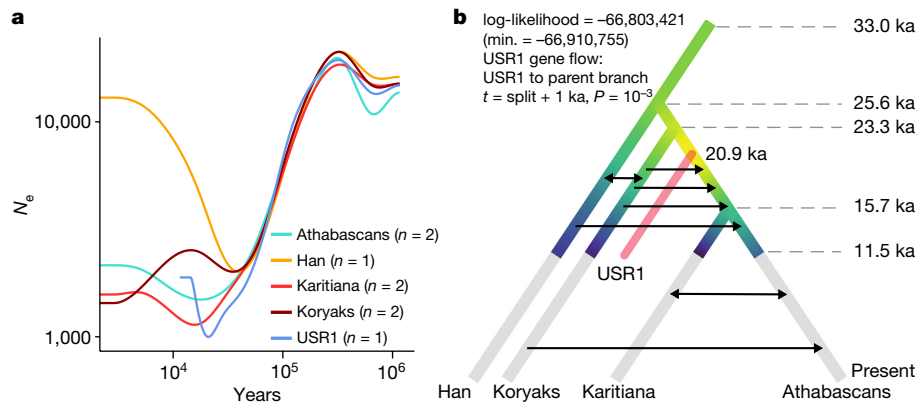


Figure 4 | USR1 demographic history in the context of East Asians, Siberians and other Native Americans. **a**, SMC++-inferred effective population sizes with respect to time for Athabascans (NNA), Karitiana (SNA), Han, Koryaks and USR1 (Supplementary Information section 19.1). We used these demographic histories as a basis for fitting a joint model for these populations. **b**, A ‘backbone demography’ was fitted excluding USR1 using momi2, a maximum likelihood approach based

Using frequency-based and ‘enhanced’ D statistics, we found no support for USR1 being closer to Papuans (a proxy for Australasians) than other Native Americans.

We leveraged the position of USR1 on the Native American branch before the NNA–SNA split to re-assess the origins of Athabaskan and Eskimo populations by fitting admixture graphs. We considered a whole-genome dataset, including Siberian, East Asian, Native American and Eskimo groups, as well as Mal’ta (Supplementary Information section 17). The heuristic approach in TreeMix²³ showed that the best proxies for the Asian component in Athabascans and Greenlandic Inuit are Koryaks and the Saqqaq individual, respectively. We then used an incremental approach to fit an f -statistic-based admixture graph¹⁹, including the Kets, which have previously been suggested to share a linguistic and perhaps a genetic link with Athabascans^{10,26}. This approach recapitulated the TreeMix results, and yielded a model in which both Athabascans and Greenlandic Inuit derive from the NNA branch. However, the Asian ancestry in Athabascans is most closely related to the Asian component in Koryaks, whereas the Saqqaq genome is the best proxy for the Siberian component in the Greenlandic Inuit (Fig. 3). We infer the latter is a consequence of Palaeo- and Neo-Eskimos having been derived from a similar Siberian population^{1,15}. This model appears to be a good fit to the data, as the observed f statistic that deviated the most from the model prediction yielded $Z = 3.27$. We also tested the robustness of this model and these predictions by computing individual D statistics and by re-fitting the model using alternative datasets (Supplementary Information section 17.3).

Finally, we inferred the demographic history of USR1 with respect to Native Americans, Siberians and East Asians, using two independent methods: diCal2²⁷ and momi2²⁸ (Supplementary Information sections 18, 19). diCal2 results indicate that the founding population of USR1, Native Americans and Siberians had a very weak structure from around 36 ka up to about 24.5 ka (Supplementary Table 7), which is when the ancestors of USR1 and Native Americans began to diverge substantially from Siberians. USR1 diverged from other Native Americans around 20.9 ka, with a period of ensuing moderate gene flow between them (Supplementary Tables 6 and 7), as indicated by a simulation study that showed a significant increase in likelihood when comparing a ‘clean split’ model to an ‘isolation with migration’ model (Supplementary Information section 18.4). Using momi2 and SMC++²⁹, we estimated a backbone demography in which Karitiana and Athabascans split around 15.7 ka, whereas their ancestral population split from Koryaks about 23.3 ka (Fig. 4). With momi2, we inferred the most likely branch (the population immediately ancestral to NNA and SNA) and time (around 21 ka) for the USR1 population to join the backbone demography,

on a site frequency spectrum (Supplementary Fig. 27), along with the most likely join-on point for USR1 onto the backbone demography (Supplementary Information section 19). We show the likelihood heat map for the latter; warmer colours correspond to a higher likelihood of USR1 joining at a given point. These estimates agree with those obtained using diCal2, a method based on haplotype data (Supplementary Information section 18).

while allowing for possible gene flow between USR and other populations (Fig. 4b and Supplementary Information section 19); results that are consistent with ref. 13 and the diCal2 inference.

These new findings, along with existing data, allow us to place Ancient Beringians within the broader context of the Pleistocene peopling of the Americas. The founding population of Native Americans (consisting of Ancient Beringians and NNA and SNA) began to diverge from ancestral Asians as early as around 36 ka, probably in northeast Asia, as there is no evidence of people in Beringia or northwest North America at this period. A high level of gene flow was maintained between them and other Asians until as late as around 25 ka^{2,13}. The subsequent isolation of the Native American founding population about 24 ka roughly corresponds to a decline in archaeological evidence for a human presence in Siberia³⁰. Both changes may result from the same underlying cause: the onset of harsh climatic conditions during the LGM². These findings, coupled with a divergence date of around 20.9 ka between USR1 and other Native Americans, are in agreement with the Beringian standstill model⁹ (Supplementary Information section 21). Ancient Beringians and the common ancestor of NNA and SNA began to diverge around 20.9 ka, after which gene flow ensued, although whether this only involved the latter or the already differentiated NNA and SNA branches cannot be determined owing to the shallow divergence times among groups.

These findings allow us to consider possible scenarios regarding where ancient Native American populations diverged (Fig. 2 and Supplementary Information sections 20, 21). Scenarios 3–5 require extended periods of strong population structure marking Ancient Beringians, NNA and SNA as separate groups, for which we do not see compelling genetic evidence; these can therefore be rejected. Scenarios 1 and 2 are compatible with our evidence of continuous gene flow among these groups, but differ as to the location of the Ancient Beringians versus NNA and SNA split at 20.9 ka, whether in northeast Asia (scenario 1) or eastern Beringia (scenario 2). Each has strengths and weaknesses relative to genetic and archaeological evidence: scenario 1 best fits the archaeological and palaeoecological evidence, as the earliest securely dated sites in Beringia are no older than around 15–14 ka, and the LGM cold period is unlikely to be associated with northward-expanding populations³⁰. Scenario 2 is genetically most parsimonious, given evidence of continuous gene flow between the Ancient Beringians and NNA and SNA, suggesting their geographical proximity 20.9–11.5 ka, and that all three were isolated from Asian and/or Siberian groups after about 24 ka and form a clade.

Scenarios 1 and 2 are both consistent with the NNA–SNA split at around 15 ka² having occurred in a region south of eastern Beringia.

The ice sheets were at that time still a substantial barrier to movement that would have helped to maintain separation from the Ancient Beringian population. Although members of the SNA branch have not been documented in regions that were once north of the Pleistocene glaciers^{1,18}, NNA groups (including Athabascan speakers) are present in Alaska today. Therefore, NNA are likely to be descendants of a population that moved north sometime after 11.5 ka²⁵.

The USR1 results provide direct genomic evidence that all Native Americans can be traced back to the same source population from a single Late Pleistocene founding event. Descendants of that population were present in eastern Beringia until at least 11.5 ka. By that time, however, a separate branch of Native Americans had already established itself in unglaciated North America, and diverged into the two basal groups that ultimately became the ancestors of most of the indigenous populations of the Americas.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 29 March; accepted 26 November 2017.

Published online 3 January 2018.

- Reich, D. *et al.* Reconstructing Native American population history. *Nature* **488**, 370–374 (2012).
- Raghavan, M. *et al.* Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* **349**, aab3884 (2015).
- Skoglund, P. *et al.* Genetic evidence for two founding populations of the Americas. *Nature* **525**, 104–108 (2015).
- von Cramon-Taubadel, N., Strauss, A. & Hubbe, M. Evolutionary population history of early Paleoamerican cranial morphology. *Sci. Adv.* **3**, e1602289 (2017).
- Hoffecker, J. F., Elias, S. A., O'Rourke, D. H., Scott, G. R. & Bigelow, N. H. Beringia and the global dispersal of modern humans. *Evol. Anthropol.* **25**, 64–78 (2016).
- Potter, B. A., Irish, J. D., Reuther, J. D. & McKinney, H. J. New insights into Eastern Beringian mortuary behavior: a terminal Pleistocene double infant burial at Upward Sun River. *Proc. Natl Acad. Sci. USA* **111**, 17060–17065 (2014).
- Rasmussen, M. *et al.* The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* **506**, 225–229 (2014).
- Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014).
- Tamm, E. *et al.* Beringian standstill and spread of Native American founders. *PLoS ONE* **2**, e829 (2007).
- Flegontov, P. *et al.* Na-Dene populations descend from the Paleo-Eskimo migration into America. Preprint at <https://doi.org/10.1101/074476> (2016).
- Dillehay, T. D. *et al.* Monte Verde: seaweed, food, medicine, and the peopling of South America. *Science* **320**, 784–786 (2008).
- Goebel, T. & Potter, B. A. In *The Oxford Handbook of the Prehistoric Arctic* (eds Friesen, T. M. & Mason, O. K.) 223–252 (Oxford Univ. Press, 2016).
- Llomas, B. *et al.* Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci. Adv.* **2**, e1501385 (2016).
- Rasmussen, M. *et al.* The ancestry and affiliations of Kennewick Man. *Nature* **523**, 455–458 (2015).
- Raghavan, M. *et al.* The genetic prehistory of the New World Arctic. *Science* **345**, 1255832 (2014).
- Rasmussen, M. *et al.* Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463**, 757–762 (2010).
- Tackney, J. C. *et al.* Two contemporaneous mitogenomes from terminal Pleistocene burials in eastern Beringia. *Proc. Natl Acad. Sci. USA* **112**, 13833–13838 (2015).
- Verdu, P. *et al.* Patterns of admixture and population structure in native populations of northwest North America. *PLoS Genet.* **10**, e1004530 (2014).
- Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
- Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
- Skotte, L., Korneliussen, T. S. & Albrechtsen, A. Estimating individual admixture proportions from next generation sequencing data. *Genetics* **195**, 693–702 (2013).
- Malaspina, A. S. *et al.* bammds: a tool for assessing the ancestry of low-depth whole-genome data using multidimensional scaling (MDS). *Bioinformatics* **30**, 2962–2964 (2014).
- Pickrell, J. K. & Pritchard, J. K. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967 (2012).
- Dyke, A. S., Moore, A. & Robertson, L. Deglaciation of North America. (Natural Resources Canada, 2003).
- Pedersen, M. W. *et al.* Postglacial viability and colonization in North America's ice-free corridor. *Nature* **537**, 45–49 (2016).
- Kari, J. M. & Potter, B. A. (eds) *The Dene-Yeniseian Connection* (Univ. Alaska, 2011).
- Steinrücken, M., Kamm, J. A. & Song, Y. S. Inference of complex population histories using whole-genome sequences from multiple populations. Preprint at <https://doi.org/10.1101/026591> (2015).
- Kamm, J. A., Terhorst, J. & Song, Y. S. Efficient computation of the joint sample frequency spectra for multiple populations. *J. Comput. Graph. Stat.* **26**, 182–194 (2017).
- Terhorst, J., Kamm, J. A. & Song, Y. S. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat. Genet.* **49**, 303–309 (2017).
- Goebel, T. The 'microblade adaptation' and recolonization of Siberia during the late Upper Pleistocene. *Archaeol. Pap. Am. Anthropol. Assoc.* **12**, 117–131 (2002).

Supplementary Information is available in the online version of the paper.

Acknowledgements The Upward Sun River excavations and analysis were conducted under a Memorandum of Agreement (MOA) signed by the State of Alaska, the National Science Foundation, the Healy Lake Tribal Council and the Tanana Chiefs Conference. We appreciate the cooperation of all parties. We thank M. Allentoft, S. Gopalakrishnan, T. Korneliussen, P. Librado, J. Ramos-Madruga, G. Renaud and F. Vieira for discussions, and the Danish National High-throughput Sequencing Centre for assistance with data generation. GeoGenetics members were supported by the Lundbeck Foundation and the Danish National Research Foundation (DNRF94) and KU2016. J.V.M.-M. was supported by Conacyt (Mexico). Samples were recovered during excavations by B.A.P. supported by NSF Grants 1138811 and 1223119. Research was supported in part by NIH grant R01-GM094402 (M.St., J.T., J.A.K. and Y.S.S.) and a Packard Fellowship for Science and Engineering (Y.S.S.). Y.S.S. is a Chan Zuckerberg Biohub investigator. D.J.M. is supported by the Quest Archaeological Research Fund. A.-S.M. is supported by the Swiss National Science Foundation and the ERC.

Author Contributions The project was conceived by E.W. and B.A.P. and headed by E.W. L.V. processed ancient DNA. J.V.M.-M. and S.R. assembled datasets. J.V.M.-M., M.St., J.T., J.A.K. and A.A. analysed genetic data. B.A.P. led the USR field investigation and B.A.P. and D.J.M. provided anthropological contextualization. B.A.P., J.D.R. and J.D.I. conducted archaeological and bioanthropological work. R.N., Y.S.S., M.Si., A.-S.M., and L.O. supervised bioinformatic and statistical analyses. B.A.P. engaged with indigenous communities. J.V.M.-M., B.A.P., D.J.M. and E.W. wrote the manuscript with input from L.V., A.-S.M., M.Si., R.S.M., L.O., Y.S.S., R.N. and the other authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to E.W. (ewillerslev@snm.ku.dk).

METHODS

Laboratory procedures. Ancient DNA work was conducted in dedicated clean laboratory facilities at the Centre for GeoGenetics, Natural History Museum, University of Copenhagen. We prepared bone powder from remains of the pars petrosa of both USR individuals and extracted DNA following previously published protocols³¹. Double-stranded dual-indexed Illumina libraries were built from uracil-specific excision reagent (USER)- and non-USER-treated extracts and were paired-end sequenced (2×75 bp) on Illumina HiSeq 2500 instruments (Supplementary Information section 2).

Sequence data processing. Raw reads were trimmed for Illumina adaptor sequences and overlapping pairs were collapsed into single reads using AdapterRemoval³². Collapsed reads were mapped to the human reference genome build 37 using BWA v.0.6.2-r126³³; seeding ($-I$ parameter) was disabled in order to prevent 5' terminal substitutions characteristic of ancient DNA to bias the mapping³⁴. Reads with mapping quality lower than 30 were discarded, PCR duplicates were removed using MarkDuplicates (<http://picard.sourceforge.net>) and local realignment was performed using GATK³⁵. We called USR1 genotypes using SAMtools mpileup³⁶ and applied the standard filters described in ref. 2. Called genotypes were phased with shapeit2-r727³⁷ using the 1,000 genomes phased variant panel (phase 3) as a reference and the HapMap recombination rates as a proxy for the genetic map of the human genome. Sites not included in the 1,000 genomes reference panel were kept as 'unphased' genotypes. Finally, we masked the dataset using a 35-mer 'snability' mask with a stringency of 0.5 (<http://lh3lh3.users.sourceforge.net/snability.shtml>) (Supplementary Information section 3).

Ancient DNA data authentication. We assessed the authenticity of the ancient DNA data by examining the fragment length distributions and the base substitution patterns across non-USER-treated reads using bamdamage²². We estimated mtDNA contamination using contamMix³⁸ on the basis of a majority rule mtDNA consensus sequence and an alignment of 311 worldwide mtDNA sequences³⁹. Nuclear contamination was estimated using the two-population model implemented in DICE⁴⁰, for which we used the 1,000 Genomes Project 'CEU' population as the putative contaminant and the 'YRI' population as the 'anchor'. Sequencing and genotyping error rates relative to a 'high-quality' sample were obtained following the method described in ref. 41 (Supplementary Information section 4).

Relatedness between USR individuals. We explored the familial relationship between both USR individuals by using NGSrelate⁴² and relate⁴³. Given the unavailability of allele frequency data for the Ancient Beringian population, we used allele frequencies from the 1,000 Genomes Project 'PEL' population as a proxy, which limited the resolution of these analyses (Supplementary Information section 5).

Reference datasets. We compared the genomes of the USR individuals to a set of 49 worldwide contemporary and ancient genomes and a SNP array dataset comprising 2,537 contemporary individuals from 167 ethnic groups (enriched in Native Americans), genotyped across 199,285 SNP sites. For the latter, European and African ancestry tracts were masked in Native American individuals (Supplementary Information 6).

Population structure analyses. We investigated the relationship between USR1, a set of ancient genomes and the SNP array reference dataset using multi-dimensional scaling as implemented in bammds²². Additionally, we explored the genetic ancestry components in the reference panel using ADMIXTURE²⁰. We obtained the most likely ancestry proportions in the ancient genomes on the basis of allele frequencies inferred by ADMIXTURE, through the genotype likelihood-based optimization method described in ref. 21 (Supplementary Information sections 7, 8).

f statistics. We computed f_3 statistics to measure the shared drift between by two particular populations or genomes, and used 'basic' and 'enhanced' D statistics to formally test hypotheses of treeness and gene-flow. We used admixtools¹⁹ for allele-frequency-based tests and ANGSD⁴⁴ for single genome tests. For both tools, standard errors were estimated through a weighted block jackknife approach over approximately 5-Mb blocks (Supplementary Information sections 9–13).

Admixture graph fitting using TreeMix. We used the heuristic approach in TreeMix²³ to assess the phylogenetic placement of USR1 in the broader context of Eurasian and Native American populations and to explore the origin of the Na-Dene and Inuit (see 'Admixture graph fitting using qpGraph'). We restricted the analysis to transversion sites where all considered populations have at least one individual with a non-missing genotype call. We grouped the resulting number of SNPs into approximately 5-Mb blocks to account for linkage disequilibrium, and for each number of migrations we ran 1,000 replicates with random seeds and kept the run with the highest likelihood. We estimated the support for internal nodes and migration edges through a bootstrap procedure (Supplementary Information sections 14, 17).

Pairwise branch lengths and genomic divergence. We used the method from ref. 7 to measure the amount of drift leading to different pairs of genomes after their split. We restricted this analysis to sites that are variable in five African genomes and obtained the counts for each of the five possible genotype configurations between a given pair of genomes, after which we used numerical optimization to infer maximum likelihood parameters (Supplementary Information section 15). We computed the average DNA divergence between pairs of genomes using the triangulation method from ref. 45, and estimated standard errors using a weighted block jackknife approach over 5-Mb blocks (Supplementary Information section 16).

Admixture graph fitting using qpGraph. We used a two-step approach to assess the origin of the Na-Dene and Inuit. First, we found the most likely Eurasian ancestry sources for these groups by using TreeMix. We then fitted f -statistics-based admixture graphs^{1,19} incrementally, such that for each new 'admixed leaf' we enumerated all possible pairs of edges using ref. 46 and kept the admixture event that produced the graph with the best maximum $|Z|$ and fitting scores. We assessed the robustness of this model and its predictions using pooled D statistics and by fitting the model using alternative datasets (Supplementary Information section 17).

Demographic inference using the sequentially Markov coalescent. We used diCal2²⁷ to estimate the key demographic parameters relating pairs of genomes including USR1 (sample dated to 11.5 ka) and a set of present-day Asian and Native American genomes. We analysed these pairs under different models, including a clean split, isolation with migration until the present, isolation with migration with a stopping time and isolation with migration with a stopping time and a second contact. We tested competing models through a simulation study and obtained confidence intervals for the inferred parameters through a parametric bootstrap strategy (Supplementary Information section 18).

Demographic inference using the site frequency spectrum. We used a combination of SMC++²⁹ and momi2²⁸ to infer demographic parameters for USR1 and a set of present-day genomes. We estimated the marginal sizes over time for each population using SMC++. We used these demographic histories as a basis for fitting a joint 'backbone demography' for the present-day populations using momi2. We then inferred the most likely join-on point for USR1 onto the backbone demography using momi2. Confidence intervals were obtained through a parametric bootstrap strategy (Supplementary Information section 19).

Data availability. Sequence data were deposited in the ENA under accession: PRJEB20398.

- Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
- Lindgreen, S. AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res. Notes* **5**, 337 (2012).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
- Schubert, M. *et al.* Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* **13**, 178 (2012).
- DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
- Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Delaneau, O., Zagury, J.-F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* **10**, 5–6 (2013).
- Fu, Q. *et al.* A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).
- Green, R. E. *et al.* A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* **134**, 416–426 (2008).
- Racimo, F., Renaud, G. & Slatkin, M. Joint estimation of contamination, error and demography for nuclear DNA from ancient humans. *PLoS Genet.* **12**, e1005972 (2016).
- Orlando, L. *et al.* Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**, 74–78 (2013).
- Korneliussen, T. S. & Moltke, I. NgsRelate: a software tool for estimating pairwise relatedness from next-generation sequencing data. *Bioinformatics* **31**, 4009–4011 (2015).
- Albrechtsen, A. *et al.* Relatedness mapping and tracts of relatedness for genome-wide data in the presence of linkage disequilibrium. *Genet. Epidemiol.* **33**, 266–274 (2009).
- Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).
- Green, R. E. *et al.* A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).
- Leppälä, K., Nielsen, S. V. & Mailund, T. admixturegraph: an R package for admixture graph manipulation and fitting. *Bioinformatics* **33**, 1738–1740 (2017).

Precision editing of the gut microbiota ameliorates colitis

Wenhan Zhu^{1*}, Maria G. Winter^{1*}, Mariana X. Byndloss², Luisella Spiga¹, Breck A. Duerkop³, Elizabeth R. Hughes¹, Lisa Büttner¹, Everton de Lima Romão², Cassie L. Behrendt³, Christopher A. Lopez², Luis Sifuentes-Dominguez⁴, Kayci Huff-Hardy⁵, R. Paul Wilson^{6†}, Caroline C. Gillis¹, Çağla Tükel⁶, Andrew Y. Koh^{1,4}, Ezra Burstein⁵, Lora V. Hooper^{3,7}, Andreas J. Bäuml² & Sebastian E. Winter¹

Inflammatory diseases of the gastrointestinal tract are frequently associated with dysbiosis^{1–8}, characterized by changes in gut microbial communities that include an expansion of facultative anaerobic bacteria of the Enterobacteriaceae family (phylum Proteobacteria). Here we show that a dysbiotic expansion of Enterobacteriaceae during gut inflammation could be prevented by tungstate treatment, which selectively inhibited molybdenum-cofactor-dependent microbial respiratory pathways that are operational only during episodes of inflammation. By contrast, we found that tungstate treatment caused minimal changes in the microbiota composition under homeostatic conditions. Notably, tungstate-mediated microbiota editing reduced the severity of intestinal inflammation in mouse models of colitis. We conclude that precision editing of the microbiota composition by tungstate treatment ameliorates the adverse effects of dysbiosis in the inflamed gut.

In genetically susceptible rodents, a dysbiotic microbiota is vertically transmissible; the affected offspring are more likely to develop intestinal inflammation^{3,9,10}, suggesting that components of the microbiota can instigate host responses in a disease-prone setting. The close association between mucosal inflammation and gut-microbiota dysbiosis poses a challenge to establishing causality between these two events. Using metagenomic sequencing, we recently identified molybdenum-cofactor-dependent metabolic pathways as a signature of inflammation-associated dysbiosis¹¹. Molybdenum-cofactor-dependent anaerobic respiratory enzymes and formate dehydrogenases contribute independently to the bloom of model Enterobacteriaceae such as *Escherichia coli*^{11,12}. We reasoned that identification of molybdenum-cofactor-dependent processes as drivers of dysbiosis would allow us to devise a strategy to manipulate microbiota metabolism and composition during gut inflammation. Selective editing of the microbiota would enable investigation of potential consequences of dysbiosis, such as exacerbation of mucosal inflammation.

Tungsten (W) can replace molybdenum in the molybdopterin cofactor, rendering this cofactor inactive in Enterobacteriaceae¹³. Supplementation of growth media with sodium tungstate does not have a general effect on growth of Enterobacteriaceae under standard aerobic laboratory conditions, but it abolishes anaerobic nitrate-reductase activity¹³ in commensal *E. coli*, *Proteus* spp., and *Enterobacter cloacae* (Fig. 1a–c, Extended Data Fig. 1a). To test whether tungstate supplementation could negate the fitness advantage conferred by anaerobic respiration and formate oxidation *in vitro*, we analysed anaerobic growth of wild-type *E. coli* strains (K-12 and Nissle 1917) and isogenic molybdenum-cofactor biosynthesis-deficient mutants ($\Delta moaA$) in

mucin broth supplemented with sodium tungstate (Fig. 1d, e, Extended Data Fig. 1b). In the presence of an electron acceptor such as nitrate, or an electron donor such as formate, the wild-type strains outcompeted the isogenic *moaA* mutants, but this fitness advantage was abrogated by the addition of tungstate (Fig. 1d, e, Extended Data Fig. 1b).

To investigate whether tungstate could inhibit molybdenum-cofactor-dependent processes in the mammalian gut, we used a mouse model of chemically induced colitis (dextran sulfate sodium (DSS)-induced colitis) in conjunction with experimentally introduced *E. coli* indicator strains. Groups of DSS- and mock-treated C57BL/6 mice were inoculated orally with an equal mixture of the *E. coli* K-12 wild-type strain and the $\Delta moaA$ mutant after the onset of inflammation. Colonization of the caecum and colon lumen was assessed five days after inoculation (Fig. 1f, g, Extended Data Fig. 2a). Prior to inoculation with *E. coli* K-12, we were unable to isolate any endogenous Enterobacteriaceae family members from these animals. Consistent with previous results¹², the K-12 wild-type strain outcompeted the $\Delta moaA$ mutant in the caecal and colonic content of DSS-treated mice (Fig. 1f). Administration of tungstate in the DSS-induced-colitis model abrogated the fitness advantage conferred by molybdenum-cofactor-dependent enzymes (Fig. 1f) and decreased overall numbers of *E. coli* K-12 in the gut lumen by several orders of magnitude (Fig. 1g). Similar observations were made using the human *E. coli* strain Nissle 1917 (Fig. 1h, Extended Data Fig. 3a, b) and a mouse *E. cloacae* strain (Fig. 1i, Extended Data Fig. 4a, b). Furthermore, the adherent-invasive *E. coli* (AIEC) strain NRG857c, originally isolated from a patient with inflammatory bowel disease, outcompeted the isogenic $\Delta moaA$ mutant in the intestinal content of DSS-treated mice (Fig. 1j). Tungstate administration negated the fitness advantage conferred by molybdenum-cofactor biosynthesis and reduced NRG857c colonization (Fig. 1j, Extended Data Figs 1c, 4c). Similarly, tungstate treatment decreased intestinal colonization by the mouse AIEC strain NC101 in a piroxicam-accelerated *Il10*^{−/−} mouse model of colitis (Fig. 1k). Taken together, these experiments based on bacterial model organisms indicate that orally administered tungstate inhibits the molybdenum-cofactor-dependent bloom of Enterobacteriaceae in mouse models of colitis.

Next, we investigated the effect of tungstate treatment on the microbiota. C57BL/6 mice that naturally harboured endogenous Enterobacteriaceae were treated with DSS, DSS plus tungstate, tungstate alone or mock treatment. After nine days, DNA extracted from the caecal content was analysed by shotgun-metagenomic sequencing and 16S profiling (Fig. 2, Extended Data Fig. 2b). Intestinal inflammation was accompanied by changes in the predicted coding capacity of

¹Department of Microbiology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390, USA. ²Department of Medical Microbiology and Immunology, School of Medicine, University of California, Davis, One Shields Avenue, Davis, California 95616, USA. ³Department of Immunology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390, USA. ⁴Department of Pediatrics, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390, USA. ⁵Department of Internal Medicine, Division of Digestive & Liver Diseases, University of Texas Southwestern Medical Center 75390, 5323 Harry Hines Boulevard, Dallas, Texas, USA. ⁶Department of Microbiology and Immunology, Lewis Katz School of Medicine, Temple University, 1801 North Broad Street, Philadelphia, Pennsylvania 19122, USA. ⁷Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390, USA. †Present address: GlaxoSmithKline, 1250 South Collegeville Road, Collegeville, Pennsylvania 19426, USA.

*These authors contributed equally to this work.

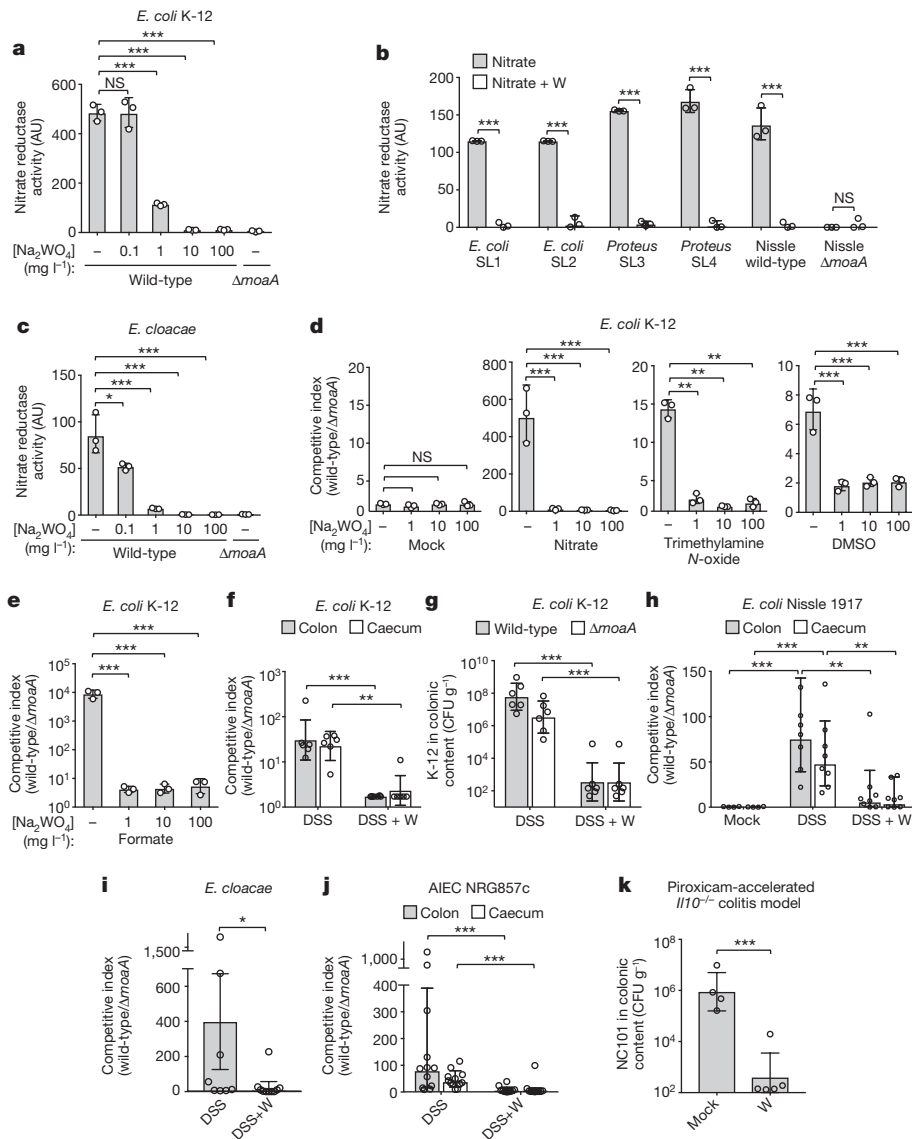


Figure 1 | Effect of tungstate on molybdenum cofactor-dependent anaerobic respiration. **a–c** Nitrate reductase activity in *E. coli* K-12 (**a**), isolated commensal Enterobacteriaceae strains SL1–SL4 and *E. coli* Nissle 1917 (**b**; strains are described in Supplementary Table 1), and an *Enterobacter cloacae* strain (**c**). W, tungstate (Na₂WO₄); AU, arbitrary units. **d, e**, Competitive anaerobic growth of the *E. coli* K-12 wild type and a *moaA* mutant (Δ*moaA*) in the presence of electron acceptors (**d**) or microaerobic growth with the electron donor formate (**e**). In **a–e**, *n* = 3 biological replicates for each condition. **f–j**, C57BL/6 mice received tungstate, DSS, or both (DSS+W) in drinking water for four days. Animals were inoculated intragastrically with an equal mixture of the indicated *E. coli* wild-type strains and isogenic *moaA* mutants. *E. coli* populations in the caecal and colonic content were analysed five days after inoculation: competitive index (**f**) and total population (**g**) of *E. coli* K-12; competitive index for Nissle 1917 (**h**), *E. cloacae* (**i**) and NRG857c (**j**). CFU, colony-forming units. In **f**, *n* = 6 per group. In **h**: Mock, *n* = 4; DSS, *n* = 8; DSS+W, *n* = 8. In **i**: DSS, *n* = 8; DSS+W, *n* = 10. In **j**: DSS, *n* = 12; DSS+W, *n* = 10. **k**, *Il10*^{-/-} mice on piroxicam-fortified diet were inoculated intragastrically with the mouse strain NC101 and received tungstate in drinking water or mock treatment. The abundance of *E. coli* NC101 was assessed after 14 days (mock, *n* = 4; W, *n* = 5). Unless stated otherwise, *n* indicates the number of animals per group. Data are shown as geometric mean and geometric s.d.; **P* < 0.05; ***P* < 0.01; ****P* < 0.001; NS, not statistically significant.

the microbiota (Fig. 2a). As found in a recent metagenomic analysis of mock and DSS-treated animals¹¹, molybdenum-cofactor-dependent processes such as nitrate respiration, trimethylamine *N*-oxide respiration and formate oxidation were overrepresented¹¹ (Fig. 2b, Extended Data Fig. 5a). Tungstate administration during colitis abolished these alterations in the metagenome (Fig. 2b, Extended Data Fig. 5a). Mirroring these changes in coding capacity, tungstate treatment during DSS-induced colitis shifted the microbial community profile from a dysbiotic state towards the normal state (Fig. 2c–e, Extended Data Fig. 5b). Consistent with the idea that molybdenum-cofactor-dependent processes contribute to the inflammation-associated bloom of *E. coli* and other Enterobacteriaceae, tungstate administration selectively blunted the expansion of the Enterobacteriaceae population, whereas other major taxonomic families were only marginally affected (Fig. 2e–g, Extended Data Fig. 5c).

In the absence of inflammation, tungstate treatment did not affect the coding capacity, diversity, community structure, or population of native Enterobacteriaceae (Fig. 2c, e, f, Extended Data Figs 5b, c, 6a). Obligate anaerobic commensals such as *Bacteroides* spp. perform a rudimentary form of anaerobic respiration by reducing endogenous fumarate to succinate. The *Bacteroides* fumarate reductase is not predicted to contain a molybdenum-cofactor-binding site and tungstate treatment had no significant effect on the prevalence of predicted fumarate-reduction pathways in the microbiome (Extended Data Fig. 6b). Furthermore, *in vivo* tungstate treatment did not affect butyrate production pathways,

a major metabolic function of the microbiota (Extended Data Fig. 6c). Supplementation of growth media with tungstate did not inhibit bacterial growth or production of succinate and butyrate by *Bacteroides* and *Clostridium* strains *in vitro* (Extended Data Fig. 6d–h). We did not observe any negative effects of tungstate on the mouse host (Extended Data Fig. 7). Collectively, these experiments support the idea that tungstate inhibits the inflammation-associated changes in gut microbiota composition that are driven by molybdenum-cofactor-dependent metabolic pathways, in particular the inflammation-associated expansion of the Enterobacteriaceae population.

We then explored the consequences of tungsten-mediated microbiota editing on mucosal inflammation in the DSS-induced-colitis model. We analysed pathological changes, colon length, mRNA levels of pro-inflammatory markers in the caecum and proximal colon, and animal body weight in mice harbouring endogenous Enterobacteriaceae in the DSS-induced-colitis model. We also analysed mice that were experimentally colonized with *E. coli* strains K-12, Nissle 1917, AIEC NRG857c or *E. cloacae*. Administration of tungstate significantly reduced inflammatory markers and pathological changes in the large intestinal mucosa, rescued the inflammation-associated reduction of colon length and ameliorated body weight loss (Fig. 3a–c, Extended Data Figs 2c–l, 3c–h, 4d–h). This was not due to reduced DSS intake during treatment (Extended Data Fig. 8a). Similarly, tungstate administration in a piroxicam-accelerated *Il10*^{-/-} colitis model reduced intestinal inflammation (Fig. 3d–g). These findings raised the possibility that

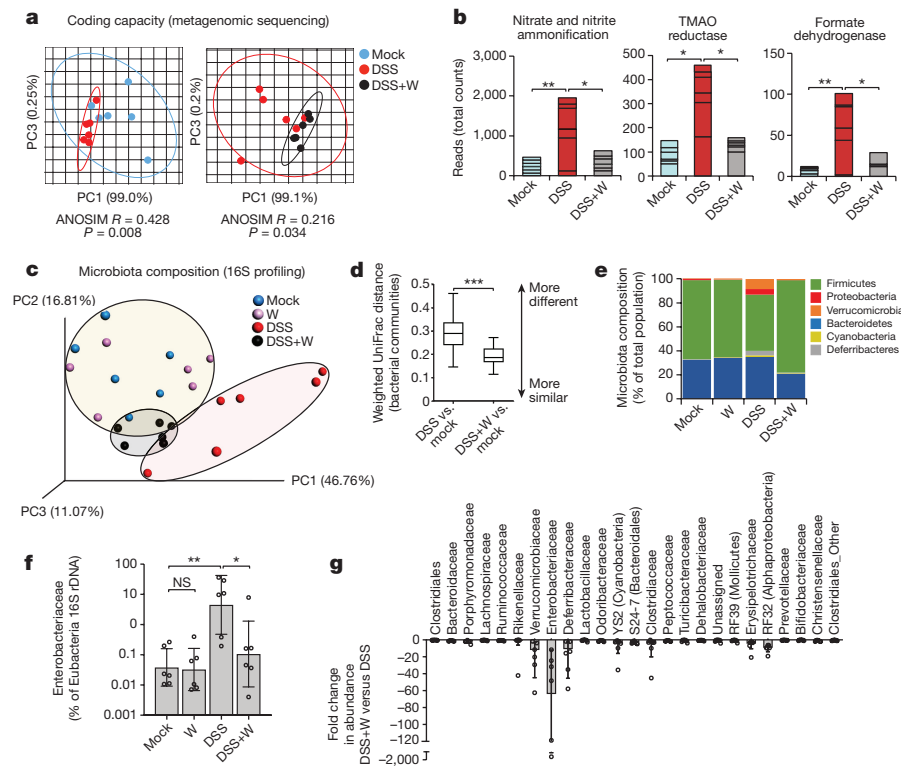


Figure 2 | Effect of tungstate treatment on composition of gut bacterial community and metabolic landscape. DNA extracted from the caecal contents of C57BL/6 mice ($n = 6$ per group) receiving the indicated treatments was analysed by metagenomic sequencing and 16S profiling. **a**, Principal coordinates analysis (PCoA) plots and analysis of similarity (ANOSIM) of the predicted coding capacity. Ellipses in **a** denote 95% confidence intervals. **b**, Tallied metagenomic reads mapped to anaerobic respiration and formate utilization pathways. **c**, PCoA of the microbiota composition (weighted UniFrac distances). **d**, Box-and-whisker plot (boxes show median, first and third quartiles, whisker denotes minimum to maximum range) of intercommunity β -diversity determined by weighted 16S UniFrac distances. **e**, Phylum-level microbiota composition. **f**, Abundance of Enterobacteriaceae quantified by qPCR. **g**, Changes in the population size of the 25 most abundant operational taxonomic units as the result of tungstate administration in the DSS-induced-colitis model. Unless otherwise noted, data are shown as geometric mean and geometric s.d.

tungsten-mediated manipulation of the gut microbiota could ameliorate gut inflammation. Alternatively, one could hypothesize that tungstate exerted anti-inflammatory effects directly on the host immune system. To test the latter hypothesis, we treated groups of germ-free C57BL/6

mice with DSS and tungstate or DSS alone for nine days and analysed the intestinal inflammatory responses. Treatment of germ-free mice with DSS resulted in moderate inflammation compared to germ-free control mice. Concomitant administration of tungstate did not

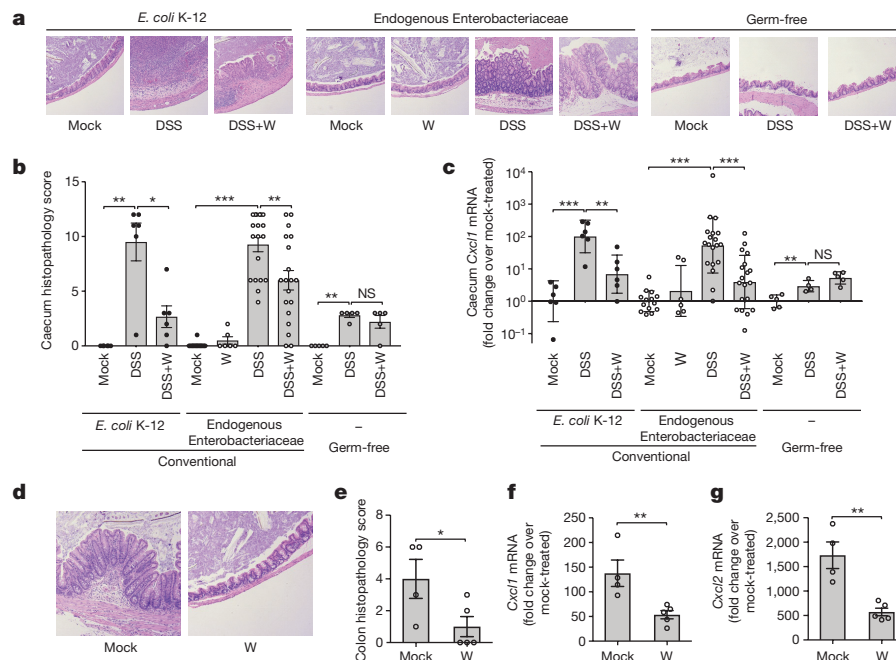


Figure 3 | Influence of tungstate treatment on mucosal inflammation. **a–c**, Conventionally raised C57BL/6 mice, treated with DSS or DSS and tungstate for four days, were inoculated with *E. coli* K-12 and samples were analysed after five days. C57BL/6 mice with a naive microbiota (including endogenous Enterobacteriaceae) or germ-free C57BL/6 mice were treated similarly with tungstate, DSS or DSS plus tungstate. *E. coli* K-12: $n = 6$ for all groups. Endogenous Enterobacteriaceae: mock, $n = 14$; W, $n = 6$; DSS, $n = 19$; DSS+W, $n = 19$. Germ-free: $n = 5$ for all groups (except in **c**; DSS, $n = 4$). **a**, Representative images of haematoxylin and eosin-stained caecal sections. **b**, Cumulative histopathology score for the caecum; data are shown

as mean and s.e.m., and each dot represents one animal. **c**, Transcription of *Cxcl1* (also known as KC) in the caecal mucosa, determined by RT-qPCR. **d–g** Groups of *Il10*^{-/-} mice were inoculated orally with *E. coli* NC101. Animals received piroxicam-fortified diet or piroxicam-fortified diet plus tungstate in drinking water for two weeks; mock, $n = 4$; W, $n = 5$. **d**, Representative images of haematoxylin and eosin-stained colonic sections. **e**, Cumulative histopathology score for the colon; data are shown as mean and s.e.m., and each dot represents one animal. **f**, **g**, Abundance of *Cxcl1* (**f**) and *Cxcl2* (**g**) mRNA in the colonic mucosa, determined by RT-qPCR. Unless otherwise noted, data are shown as geometric mean and geometric s.d.

interfere with the induction of this response, indicating that tungsten limits intestinal inflammation by manipulating the gut microbiota (Fig. 3a–c, Extended Data Fig. 2c, d). Tungstate had no observable effect on pro-inflammatory responses or cellular resistance to DSS injury in cultured cells (Extended Data Fig. 8b–d). Therapeutic administration of tungstate after the onset of inflammation was sufficient to inhibit molybdenum-cofactor-dependent processes in *E. coli* Nissle 1917 (Extended Data Fig. 8e, f), supporting the hypothesis that the effect of tungsten on microbial populations was not due to tungstate interfering with the induction of DSS-induced inflammation. Collectively, these data suggest that tungsten limits gut inflammation through manipulation of the mouse gut microbiota.

A subset of people with inflammatory bowel disease exhibit changes in the composition of their gut microbiota that include increased abundance of Enterobacteriaceae family members¹. We humanized the gut of germ-free mice with gut microbiota from patients with active flares. To model intestinal inflammation, groups of mice were treated with either DSS alone or DSS and tungstate, and housed separately. Administration of tungstate reduced the intestinal Enterobacteriaceae load and decreased markers of mucosal inflammation (Extended Data Fig. 4i–m), thus providing evidence that the effect of tungsten is not unique to mouse microbiota.

An imbalance in the gut-associated microbial community may underlie many human diseases, but current approaches to treating dysbiosis lack the sophistication needed to restore a balanced community *in situ*. Administration of antibiotics broadly reduces numbers of many members of the gut microbiota without discriminating between beneficial and potentially harmful microbes. In some instances, removal of potentially harmful members of the community can lead to a beneficial outcome^{14–16}. However, removal of beneficial microbes can lead to pathogen expansion^{2,17} or increased bowel irritability¹⁸, thereby adversely affecting the host. Commensal Enterobacteriaceae contribute to resistance to colonization by enteric pathogens by competing for critical nutrients^{19,20}. Oral administration of probiotic *E. coli* Nissle 1917 is effective in maintaining remission in patients with ulcerative colitis²¹, and microcins produced by *E. coli* Nissle 1917 suppress the growth of pathogenic bacteria²². Thus, it might be preferable to control the population size of commensal Enterobacteriaceae in the gut microbiome than to remove them entirely. In contrast to broad-spectrum antibiotics, tungstate treatment of the dysbiotic microbiota allows selective control of bacterial populations, such as Enterobacteriaceae, that rely on molybdenum-cofactor-dependent processes. Because these molybdenum-cofactor-dependent processes operate only during gut inflammation¹¹, tungsten treatment acts only on the enterobacterial population in the disease state, and does not eliminate Enterobacteriaceae from the ecosystem during homeostatic conditions. Our work identifies molybdenum-cofactor-dependent processes as a target for controlling disease-specific aspects of the microbiota composition. Furthermore, our results provide experimental evidence that this rationally designed microbiome editing approach can improve dysbiosis-associated mucosal inflammation.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 November 2015; accepted 24 November 2017.

Published online 3 January 2018.

- Frank, D. N. *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl Acad. Sci. USA* **104**, 13780–13785 (2007).
- Lupp, C. *et al.* Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. *Cell Host Microbe* **2**, 119–129 (2007).
- Garrett, W. S. *et al.* Enterobacteriaceae act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell Host Microbe* **8**, 292–300 (2010).
- Mshvildadze, M. *et al.* Intestinal microbial ecology in premature infants assessed with non-culture-based techniques. *J. Pediatr.* **156**, 20–25 (2010).

- Vujkovic-Cvijin, I. *et al.* Dysbiosis of the gut microbiota is associated with HIV disease progression and tryptophan catabolism. *Sci. Transl. Med.* **5**, 193ra91 (2013).
- Raetz, M. *et al.* Parasite-induced TH1 cells and intestinal dysbiosis cooperate in IFN- γ -dependent elimination of Paneth cells. *Nat. Immunol.* **14**, 136–142 (2013).
- Winter, S. E., Lopez, C. A. & Bäuml, A. J. The dynamics of gut-associated microbial communities during inflammation. *EMBO Rep.* **14**, 319–327 (2013).
- Shin, N. R., Whon, T. W. & Bae, J. W. Proteobacteria: microbial signature of dysbiosis in gut microbiota. *Trends Biotechnol.* **33**, 496–503 (2015).
- Garrett, W. S. *et al.* Communicable ulcerative colitis induced by T-bet deficiency in the innate immune system. *Cell* **131**, 33–45 (2007).
- Elinav, E. *et al.* NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. *Cell* **145**, 745–757 (2011).
- Hughes, E. R. *et al.* Microbial respiration and formate oxidation as metabolic signatures of inflammation-associated dysbiosis. *Cell Host Microbe* **21**, 208–219 (2017).
- Winter, S. E. *et al.* Host-derived nitrate boosts growth of *E. coli* in the inflamed gut. *Science* **339**, 708–711 (2013).
- Gates, A. J. *et al.* Properties of the periplasmic nitrate reductases from *Paracoccus pantotrophus* and *Escherichia coli* after growth in tungsten-supplemented media. *FEMS Microbiol. Lett.* **220**, 261–269 (2003).
- Hans, W., Schölmerich, J., Gross, V. & Falk, W. The role of the resident intestinal flora in acute and chronic dextran sulfate sodium-induced colitis in mice. *Eur. J. Gastroenterol. Hepatol.* **12**, 267–273 (2000).
- Rutgeerts, P. *et al.* Controlled trial of metronidazole treatment for prevention of Crohn's recurrence after ileal resection. *Gastroenterology* **108**, 1617–1621 (1995).
- Prantera, C. *et al.* An antibiotic regimen for the treatment of active Crohn's disease: a randomized, controlled clinical trial of metronidazole plus ciprofloxacin. *Am. J. Gastroenterol.* **91**, 328–332 (1996).
- Stecher, B. *et al.* *Salmonella enterica* serovar Typhimurium exploits inflammation to compete with the intestinal microbiota. *PLoS Biol.* **5**, 2177–2189 (2007).
- Collins, S., Verdu, E., Denou, E. & Bercik, P. The role of pathogenic microbes and commensal bacteria in irritable bowel syndrome. *Dig. Dis.* **27** (Suppl 1), 85–89 (2009).
- Kamada, N. *et al.* Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. *Science* **336**, 1325–1329 (2012).
- Brugiroux, S. *et al.* Genome-guided design of a defined mouse microbiota that confers colonization resistance against *Salmonella enterica* serovar Typhimurium. *Nat. Microbiol.* **2**, 16215 (2016).
- Kruis, W. *et al.* Maintaining remission of ulcerative colitis with the probiotic *Escherichia coli* Nissle 1917 is as effective as with standard mesalazine. *Gut* **53**, 1617–1623 (2004).
- Sassone-Corsi, M. *et al.* Microcins mediate competition among Enterobacteriaceae in the inflamed gut. *Nature* **540**, 280–283 (2016).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by NIH grants AI112445 (A.J.B.), AI118807 (S.E.W.), AI128151 (S.E.W.), DK070855 (L.V.H.), DK102436 (B.A.D.) and 5K12HD-068369 (L.S.-D.), Welch Foundation grants I-1858 (S.E.W.) and I-1874 (L.V.H.), American Cancer Society Research Scholar Grant MPC-130347 (S.E.W.) and a Crohn's and Colitis Foundation of America postdoctoral fellowship no. 454921 (W.Z.). Work in the L.V.H. laboratory is supported by the Howard Hughes Medical Institute. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funders. We thank B. Sartor for the *E. coli* NC101 strain.

Author Contributions W.Z., M.G.W., L.S., L.B., E.d.L.R., R.P.W., E.R.H., C.A.L. and C.C.G. performed and analysed nitrate reductase activity, *in vitro* bacterial competitive growth, NF- κ B induction, DSS cytotoxicity experiments and experiments involving conventionally raised C57BL/6 mice. M.G.W., L.S. and E.R.H. performed *Il10*^{-/-} mouse experiments. W.Z. and M.G.W. performed inflammation analysis. C.L.B., M.G.W., L.S. and W.Z. performed germ-free-mouse experiments. B.A.D. and W.Z. analysed 16S and metagenomic data. M.X.B. analysed the histopathology. L.S.-D. and K.H.-H. contributed to humanized mouse experiments. L.S. and S.E.W. performed metabolite quantification. W.Z., C.T., A.Y.K., E.B., L.V.H., A.J.B. and S.E.W. designed the experiments, interpreted the data and wrote the manuscript with contributions from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to A.J.B. (ajbaumler@ucdavis.edu) or S.E.W. (Sebastian.Winter@UTSouthwestern.edu).

Reviewer Information Nature thanks C. Elson, M. Fischbach and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Bacterial strains. The *E. coli*, *Proteus*, and *E. cloacae* strains used in this study are listed in Supplementary Table 1. All strains were routinely grown aerobically in LB broth (10 g l⁻¹ tryptone, 5 g l⁻¹ yeast extract, 10 g l⁻¹ NaCl) or on LB agar plates (10 g l⁻¹ tryptone, 5 g l⁻¹ yeast extract, 10 g l⁻¹ NaCl, 15 g l⁻¹ agar) at 37 °C. When appropriate, antibiotics were added to the medium at the following concentrations: 30 µg ml⁻¹ chloramphenicol, 100 µg ml⁻¹ carbenicillin, 50 µg ml⁻¹ kanamycin.

Plasmids. All primers and plasmids are listed in Supplementary Tables 2 and 3. pWZ5 was constructed with standard molecular cloning techniques²³ using the Gibson Assembly Cloning Kit (New England Biolab) according to the recommendations of the manufacturer. The flanking regions of the *moaA* gene from the *E. coli* strain NRG857c were amplified and ligated into pGP706 to make pWZ5. Plasmid inserts were verified by Sanger sequencing.

Construction of mutants by allelic exchange. pWZ5 was propagated in DH5α λpir and conjugated into the *E. coli* strains NRG857c or NC101 using S17-1 λpir as the conjugative donor strain. Exconjugants that had the suicide plasmid integrated into the recipient chromosome (single crossover) were recovered on LB plates containing appropriate antibiotics. Sucrose plates (8 g l⁻¹ nutrient broth base, 5% sucrose, 15 g l⁻¹ agar) were used to select for the second crossover event, thus creating WZ12 and WZ245, respectively. Deletion of the target gene was confirmed by PCR.

Anaerobic growth assays. Anaerobic growth assays were performed in mucin broth. Mucin broth contained hog mucin (Sigma–Aldrich) at a final concentration of 0.5% (w/v) in no-carbon E medium²⁴ and was supplemented with trace elements²⁵. Sodium formate, sodium nitrate, DMSO and trimethylamine-N-oxide (TMAO, Sigma–Aldrich) were added to a final concentration of 40 mM, in the absence or presence of sodium tungstate (Sigma–Aldrich) at the indicated final concentrations. A volume of 2 ml of mucin broth was inoculated with the indicated strains at a concentration of 1 × 10⁴ colony-forming units (CFU) per ml and incubated anaerobically (Bactron EZ Anaerobic Chamber, Sheldon Manufacturing) for 18 h at 37 °C. Bacterial numbers were counted as described¹².

DSS-induced colitis model and sodium tungstate treatment. All experiments involving mice were approved by the Institutional Animal Care and Use Committee at UT Southwestern Medical Center (APN#T-2013-0159) and UC Davis (APN#16196). Studies involving animals were performed with compliance to all relevant ethical regulations. Female 9–12-week-old C57BL/6J wild-type mice were obtained from Jackson Laboratory (Bar Harbour) and bred at UT Southwestern (essentially devoid of endogenous Enterobacteriaceae) or Charles River Laboratories (Morrisville) (harbouring endogenous Enterobacteriaceae), as indicated. Mice were randomly assigned into cages before the experiment. The drinking water was replaced with either filter-sterilized water (mock treatment), a filter-sterilized solution of 0.2% (w/v) sodium tungstate (Sigma), a filter-sterilized solution of 2% or 3% (w/v) DSS (relative molecular mass 36,000–50,000; MP Biomedicals) in water, or a filter-sterilized solution of DSS and 0.2% (w/v) sodium tungstate. In one experiment, tungsten was administered in a sodium tungstate-fortified diet (1,000 parts per million (p.p.m.)). At the indicated time points, animals were inoculated orally with either 0.1 ml LB broth or 0.1 ml LB broth containing 1 × 10⁹ CFU *E. coli*, or remained uninfected. In the competitive colonization experiments, mice were inoculated with 5 × 10⁸ CFU of each *E. coli* or *E. cloacae* strain. One day before the end of the experiment, the drinking water was switched to regular, filter-sterilized water for 24 h to reduce the amount of DSS present in the samples. After euthanization, colonic and caecal tissue were collected, flash frozen and stored at –80 °C for subsequent mRNA and protein expression analysis. Faecal material, caecal content, and colonic content were collected in sterile PBS and the bacterial loads for the *E. coli* strains or Enterobacteriaceae were quantified by plating serial tenfold dilutions on LB plates supplemented with appropriate antibiotics or MacConkey agar plates, respectively. *E. coli* NC101 and Nissle 1917 strains were differentially marked with the low-copy number plasmids pWSK29 and pWSK129 to facilitate bacterial recovery from biological samples¹². For the competitive colonization experiments involving NRG857c, animals were inoculated with an equal mixture of the NRG857c Δ*lacZ* mutant (LB33) and the Δ*moaA* mutant (WZ12) as described above. The bacterial load in the luminal content of the indicated organs was determined by plating serial tenfold dilutions on LB plates supplemented with the appropriate antibiotics and 40 mg l⁻¹ 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside. Germ-free C57BL/6 mice were maintained in plastic gnotobiotic isolators on a 12-h light cycle. DSS-mediated colitis was induced in 8–12-week-old germ-free mice, following the protocol described above.

Piroxicam-accelerated colitis model in conventional *Il10*^{-/-} mice. Conventional *Il10*^{-/-} mice (7–12 weeks old, males only) on a C57BL/6 background were randomly assigned into cages before oral inoculation with 1 × 10⁹ CFU of mouse AIEC NC101. Regular mouse chow was replaced with piroxicam-fortified diet (100 p.p.m.; Teklad custom research diets, Envigo) and changed daily. Drinking

water was replaced with either filter-sterilized water (mock treatment) or a filter-sterilized solution of 0.2% (w/v) sodium tungstate. After 14 days, mice were euthanized and the samples were collected as described above.

Faecal transplant into gnotobiotic mice. All procedures involving human subjects were reviewed and approved by the institutional review board at the University of Texas Southwestern Medical Center (IRB#112010-130). Studies involving human samples were performed with compliance to all relevant ethical regulations. Written informed consent was obtained from all participants or parents or legal guardians of participating minors. Except for E.B. and S.F.-D., none of the investigators handling the samples had access to personally identifiable information. Patients were considered for faecal donation if they had an established diagnosis of inflammatory bowel disease, had active disease at the time of collection and were free from antibiotic use over the past three months. Characteristics of patients from whom samples were taken are summarized in Supplementary Table 4. Human faecal samples were obtained during colonoscopy by direct endoscopic aspiration of faecal contents from patients with active colonic disease. A total of 10 ml of liquid faecal material was collected from each patient and aliquoted into 1-ml cryovials. The samples were then snap-frozen in liquid nitrogen and stored at –80 °C until use.

Germ-free Swiss–Webster mice (7–12 weeks old, mix of male and female) were maintained in plastic gnotobiotic isolators on a 12-h light cycle. Mice were randomized, paired and orally gavaged with endoscopy samples from the patients listed in the table at the end of this section. Colonization was allowed to proceed for three days before mice received DSS or DSS plus sodium tungstate for seven days. Mice were euthanized and the samples were collected as described above.

16S RNA pyrosequencing and analysis. Caecal contents were collected and DNA was extracted from faecal samples using the MoBio PowerFecal kit (MoBio Laboratories) according to the recommendations of the manufacturer. The extracted DNA was subjected to KCl precipitation to remove residual DSS contaminants. In brief, DNA was incubated with excess KCl on ice to precipitate DSS. The samples were then cleared by centrifugation and the resulting supernatant was subsequently subjected to ethanol precipitation to recover the DNA. The purified DNA was subjected to paired end library construction to facilitate assemblies and longer accurate reads. The 16S rRNA coding sequences used to identify the bacteria were amplified using primers 515F and 806R that flank the V3–V4 hypervariable region, and barcoded before pyrosequencing. The bar-coded amplicons were purified and quantified on an Invitrogen Qubit system (Life technology). Libraries were sequenced using an Illumina MiSeq system (Illumina). 16S-sequencing data was subjected to a standard workflow for processing and quality assessment of the raw 16S-sequence data and the downstream phylogenetic analysis. The pipeline consists of an initial customized Linux-based command script for trimming, demultiplexing and quality filtering the raw paired end-sequence data generated by the Illumina system. Sequence alignment, operational taxonomic units (OTUs) picking against the Greengenes reference collection, clustering, phylogenetic and taxonomic profiling, permutational multivariate analysis of variance (PERMANOVA), and the analysis of beta diversity (principle component analysis) on the demultiplexed sequences were performed with the Quantitative Insights into Microbial Ecology (QIIME) open source software package²⁶.

Metagenomics. Groups of randomized Charles River C57BL/6 mice were treated as described in Extended Data Fig. 2b. Sample collection, shotgun metagenomics sequencing and data analysis were performed as previously described¹¹.

Reads that mapped to the SEED database were exported from MEGAN5 into BIOM tables, which were subjected to analysis of similarity (ANOSIM) in Qiime²⁶ and principal component analysis (PCA) using STAMP²⁷. To map reads to bacterial metabolic genes, a total of 100 of each of the butyrate production operons (*bcdAB*, *but* and *ato*) and succinate dehydrogenase operons (*sdhABC*) were downloaded from the KEGG database. Sequences were clustered to remove redundancy using cdhit-est^{28,29} with a sequence identity threshold of 0.9. Paired end reads were mapped to these gene clusters using the BBmap tool with the following settings: 'qtrim = lr, minid = 0.90, ambig = random, covstats = true'. Coverage statistics for each gene cluster were tallied from the percentage of unambiguous and ambiguous mapped reads and used to determine the absolute number of reads that mapped to a particular gene set. A similar strategy was used to map reads to fumarate respiration and butyrate production pathways.

Abundance of Enterobacteriaceae. The relative abundance of endogenous Enterobacteriaceae as part of the bacterial microbiota was analysed as described previously^{30–32}. In brief, the caecum or colon content was extracted using the PowerFecal DNA Isolation Kit (MoBio Laboratories) according to the manufacturer's instructions, and the resulting DNA was further purified using the KCl method as described above. A 2-µl sample of the bacterial DNA was used as the template for SYBR Green-based real-time PCR reactions as described above. The gene copy number in the sample was determined based on a standard curve generated using

pSW321 and pSW196 as previously described³³. The primers used are listed in Supplementary Table 2. The fraction of Enterobacteriaceae as part of the entire bacterial population for each sample was calculated by dividing the gene-copy number of the *Enterobacteriaceae* by the gene-copy number determined using the eubacterial primers.

Quantification of mRNA levels in intestinal tissue. The relative transcription levels of mRNAs for iNOS, CXCL1, CXCL2, IL-17, IL-6, IFN- γ , LCN2 and TNF- α , encoded by the *Nos2*, *Cxcl1*, *Cxcl2*, *Il17*, *Il6*, *Ifng*, *Lcn2* and *Tnf* genes, respectively, were determined by qRT-PCR as described previously³¹. In brief, colonic or caecal tissue was homogenized in a Mini Beadbeater (Biospec Products) and RNA was extracted using the TRI-reagent method (Molecular Research Center). To remove residual DSS contaminants, RNA was further purified using the Dynabeads mRNA Direct Kit (Life Technologies) per the manufacturer's instructions. cDNA was generated with TaqMan reverse-transcription reagents (Life Technologies). Real-time PCR was performed using SYBR Green (Life Technologies), and data were acquired in a QuantStudio 6 Flex instrument (Life Technologies) and analysed using the comparative C_t method. The primers listed in Supplementary Table 2 were added at a final concentration of 250 nM. Target-gene transcription of each sample was normalized to the respective levels of *Gapdh* mRNA.

Histopathology. Mouse caecal and colonic tissue was fixed in phosphate-buffered formalin and 5- μ m sections of the tissue were stained with haematoxylin and eosin. The fixed and stained sections were blinded and evaluated by an experienced veterinary pathologist according to the criteria described previously¹². Images were taken at a magnification of 10 \times , and the contrast for the images was uniformly (linear) adjusted using Adobe Photoshop CC.

Measurement of succinate and butyrate concentrations in bacterial culture using GC-MS. Bacterial cultures were cleared by centrifugation at 13,200g at 4°C for 30 min and then passed through a 0.22- μ m filter. The supernatant was dried using a SpeedVac concentrator. The pellet was then dissolved in pyridine at 80°C for 20 min before derivatization with *n*-tert-butyltrimethylsilyl-*n*-methyltrifluoroacetamide with 1% t-BDMCS silylation reagent (Cerilliant) at 80°C for 1 h. Derivatized samples were transferred to autosampler vials for gas chromatography-mass spectrometry (GC-MS) analysis (Shimadzu, TQ8040). The injection temperature was 250°C and the injection split ratio was set to 1:100 or 1:1,000 with an injection volume of 1 μ L. The gas chromatography oven temperature started at 130°C for 4 min, rising to 230°C at 4°C min⁻¹, and to 280°C at 20°C min⁻¹ with a final hold at this temperature for 2 min. The gas chromatography flow rate of the helium carrier gas was kept constant at a linear velocity of 50 cm s⁻¹. The column used was a 30 m \times 0.25 mm \times 0.25 μ m Rtx-5Sil MS (Shimadzu). The interface temperature was 300°C. The electron-impact ion-source temperature was 200°C, with 70 V ionization voltage and 150 μ A current. To measure succinate, Q3 scans (range of 50–500 *m/z*, 1000 *m/z* per second) were first performed to determine the retention time for succinate and succinate-2,2,3,3-d₄ (CDN Isotopes), which was 11.0 and 10.9 min respectively. Multiple-reaction-monitoring mode was then used (target ion *m/z* 289 \rightarrow 147, reference ion *m/z* 331 \rightarrow 189) to measure succinate quantitatively. To measure butyrate, Q3 scans were performed as described above, and the retention time for butyrate and butyrate-d₇ (CDN Isotopes) was 6.1 and 6.2 min, respectively. Q3-selected ion monitoring (single-quadrupole mode) with an event time of 0.05 s was performed to quantitatively measure butyrate. The target and reference (qualifier) ions for butyrate were *m/z* = 145 and *m/z* = 75, respectively; target and reference ions for deuterated butyrate were *m/z* = 152 and *m/z* = 76.

Strain isolation and identification. Tenfold serial dilutions of the faecal content of C57BL/6 mice (Charles River) were plated on MacConkey agar (10 g l⁻¹ pancreatic digest of gelatin, 3 g l⁻¹ peptone, 10 g l⁻¹ lactose, 1.5 g l⁻¹ bile salts, 5 g l⁻¹ sodium chloride, 13.5 g l⁻¹ agar, 30 mg l⁻¹ neutral red, 1 mg l⁻¹ crystal violet) and incubated aerobically at 37°C overnight. To isolate mouse-commensal *E. coli* and *Proteus* strains, single colonies were isolated and identified using the Enteropluri Test (Liofilchem) per the manufacturer's recommendations.

Nitrate reductase activity assays. Overnight cultures of *E. coli* or *Proteus* strains were diluted 1:100 in fresh LB broth containing 40 mM sodium nitrate to induce the expression of nitrate reductases, in the presence or absence of sodium tungstate at the indicated concentrations. Cultures were incubated aerobically for 3 h at 37°C and the relative nitrate reductase activity was measured as described previously³⁴. The experiment was repeated three times and representative results are shown.

NF- κ B activation in epithelial cells. HeLa57A cells, stably transfected with an NF- κ B-luciferase reporter construct^{35,36}, were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum at 37°C in a 5% CO₂ atmosphere. For the NF- κ B activation assays, cells were seeded in a 48-well plate to reach 80% confluency within 24 h. Cells were treated with 0.1, 1 or 10 ng ml⁻¹ of phorbol 12-myristate 13-acetate (PMA, dissolved in DMSO) or DMSO alone. At the same time, sodium tungstate with a final concentration of 0.02% or 0.002% (w/v) in water was added to the cells. After 5 h, cells were washed in DPBS and

lysed in 0.1 ml of reporter lysis buffer (Promega). Firefly luciferase activity was measured with a commercial luciferase assay system (Promega). The experiment was repeated three times and representative results are shown. HeLa57A cells were generated by R. T. Hay (University of Dundee). These cells have not been authenticated or tested for mycoplasma contamination.

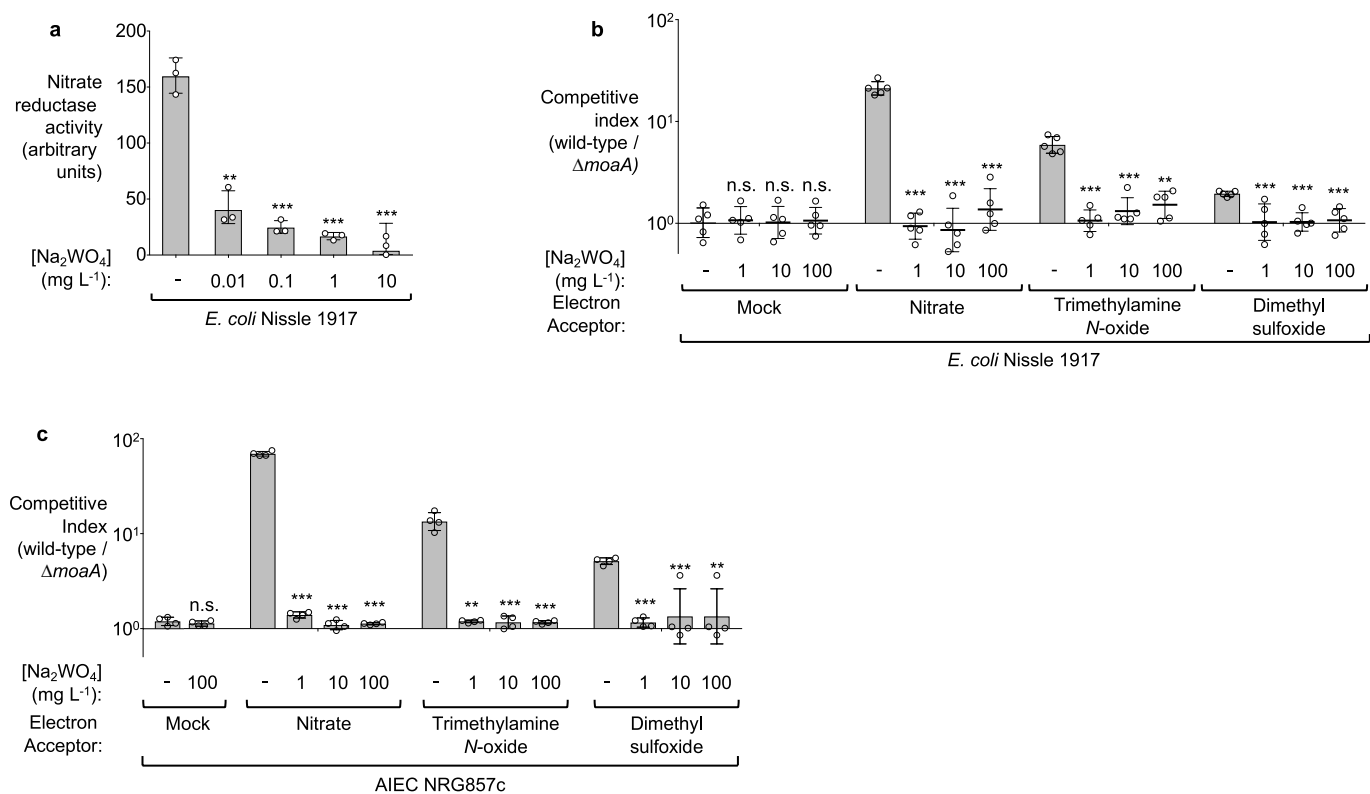
LDH-release assay. The MODE-K cell line was maintained in DMEM (Sigma) supplemented with 10% FBS at 37°C in 5% CO₂. Bone-marrow-derived macrophages (BMDMs) were differentiated from bone-marrow cells collected from femurs and tibias of SPF C57BL/6 mice. In brief, bone-marrow cells were collected with 10 ml cold RPMI-1640 medium (Sigma) and then pelleted at 1,000 r.p.m. for 5 min at 25°C. The cells were resuspended in BMM medium (RPMI-1640 supplemented with 10% heat-inactivated FBS, 1 mM glutamine, 1% antibiotics-antimycotics, and 30% L-cell conditioned medium) and allowed to differentiate for seven days. MODE-K experiments were performed in triplicate. Plates were seeded with cells to a final confluency of 80% before treatment. Two days after seeding, MODE-K cells were treated for 24 h with 2–4% DSS (Alfa Aesar), with and without 0.002–0.2% sodium tungstate dihydrate (Sigma). For BMDM experiments, plates were seeded with 1 \times 10⁵ cells per well for 48 h. After 24 h the medium was replaced with RPMI supplemented with 2% FBS and glutamine. On the day of the experiment, culture medium was replaced with medium supplemented with 4 or 6% DSS, with and without 0.2% sodium tungstate dehydrate, for 24 h. Cytotoxicity was determined using the LDH-release assay CytoTox 96 non-radioactive cytotoxicity assay (Promega), per the manufacturer's recommendations. Absorbance readings were corrected based on the absorbance of the medium alone. Five-minute treatment with 10% Triton X-100 was used as the total LDH-release control. The experiment was repeated three times, and representative results are shown. MODE-K cells were generated by D. Kaiserlian (Institut Pasteur de Lyon). These cells have not been authenticated or tested for mycoplasma contamination.

Statistics and reproducibility. No statistical methods were used to pre-determine sample size. The investigators were not blinded to allocation during experiments and outcome assessment, except for histology analysis. Nitrate reductase activities, fold changes in mRNA levels, competitive indices, relative abundance of Enterobacteriaceae and bacterial numbers were transformed logarithmically and the statistical significance of differences between groups was determined using a two-sided Student's *t*-test or PERMANOVA (using distance matrices). Cumulative histopathology scores were analysed using the Mann-Whitney *U* test. Details regarding the statistics of each experiment are reported in Supplementary Table 5.

Data availability. The bacterial 16S-ribosomal DNA and metagenomics-sequencing reads generated and analysed during the current study are available at the European Bioinformatics Institute repository under accession numbers PRJEB15095 and PRJEB19192. All data generated or analysed during this study are included in this published article (and its Supplementary Information files).

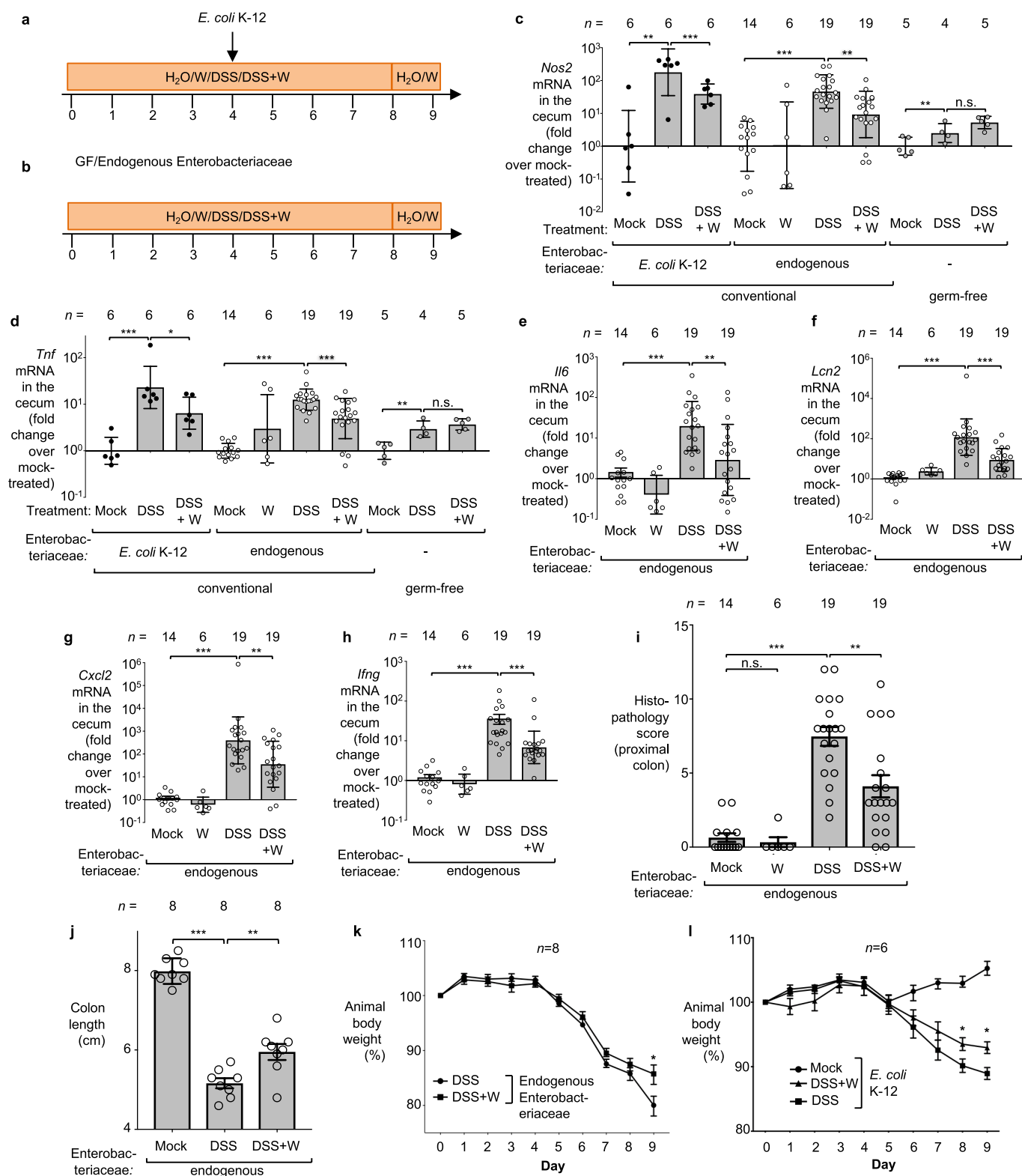
23. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning* 2nd edn (Cold Spring Harbor Laboratory Press, 1989).
24. Davis, R. W., Roth, J. R. & Botstein, D. *Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory Press, 1980).
25. Price-Carter, M., Tingey, J., Bobik, T. A. & Roth, J. R. The alternative electron acceptor tetrathionate supports B12-dependent anaerobic growth of *Salmonella enterica* serovar Typhimurium on ethanolaniline or 1,2-propanediol. *J. Bacteriol.* **183**, 2463–2475 (2001).
26. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
27. Shavit, Y., Hamey, F. K. & Lio, P. FishHiCal: an R package for iterative FISH-based calibration of Hi-C data. *Bioinformatics* **30**, 3120–3122 (2014).
28. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
29. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
30. Barman, M. et al. Enteric salmonellosis disrupts the microbial ecology of the murine gastrointestinal tract. *Infect. Immun.* **76**, 907–915 (2008).
31. Winter, S. E. et al. Contribution of flagellin pattern recognition to intestinal inflammation during *Salmonella enterica* serotype Typhimurium infection. *Infect. Immun.* **77**, 1904–1916 (2009).
32. Winter, S. E. et al. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. *Nature* **467**, 426–429 (2010).
33. Bacchetti De Gregoris, T., Aldred, N., Clare, A. S. & Burgess, J. G. Improvement of phylum- and class-specific primers for real-time PCR quantification of bacterial taxa. *J. Microbiol. Methods* **86**, 351–356 (2011).
34. Stewart, V. & Parales, J. Jr. Identification and expression of genes *narL* and *narX* of the *nar* (nitrate reductase) locus in *Escherichia coli* K-12. *J. Bacteriol.* **170**, 1589–1597 (1988).
35. Rodriguez, M. S., Thompson, J., Hay, R. T. & Dargemont, C. Nuclear retention of I κ B α protects it from signal-induced degradation and inhibits nuclear factor κ B transcriptional activation. *J. Biol. Chem.* **274**, 9108–9115 (1999).
36. Keestra, A. M. et al. A *Salmonella* virulence factor activates the NOD1/NOD2 signaling pathway. *MBio* **20**, e00266–11 (2011).

37. Eaves-Pyles, T. *et al.* *Escherichia coli* isolated from a Crohn's disease patient adheres, invades, and induces inflammatory responses in polarized intestinal epithelial cells. *Int. J. Med. Microbiol.* **298**, 397–409 (2008).
38. Kim, S. C. *et al.* Variable phenotypes of enterocolitis in interleukin 10-deficient mice monoassociated with two different commensal bacteria. *Gastroenterology* **128**, 891–906 (2005).
39. Pal, D., Venkova-Canova, T., Srivastava, P. & Chattoraj, D. K. Multipartite regulation of *rctB*, the replication initiator gene of *Vibrio cholerae* chromosome II. *J. Bacteriol.* **187**, 7167–7175 (2005).
40. Simon, R., Priefer, U. & Puhler, A. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat. Biotechnol.* **1**, 784–791 (1983).
41. Overbergh, L. *et al.* The use of real-time reverse transcriptase PCR for the quantification of cytokine gene expression. *J. Biomol. Tech.* **14**, 33–43 (2003).
42. Godinez, I. *et al.* T cells help to amplify inflammatory responses induced by *Salmonella enterica* serotype Typhimurium in the intestinal mucosa. *Infect. Immun.* **76**, 2008–2017 (2008).
43. Wilson, R. P. *et al.* The Vi-capsule prevents Toll-like receptor 4 recognition of *Salmonella*. *Cell. Microbiol.* **10**, 876–890 (2008).
44. Wang, R. F. & Kushner, S. R. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* **100**, 195–199 (1991).



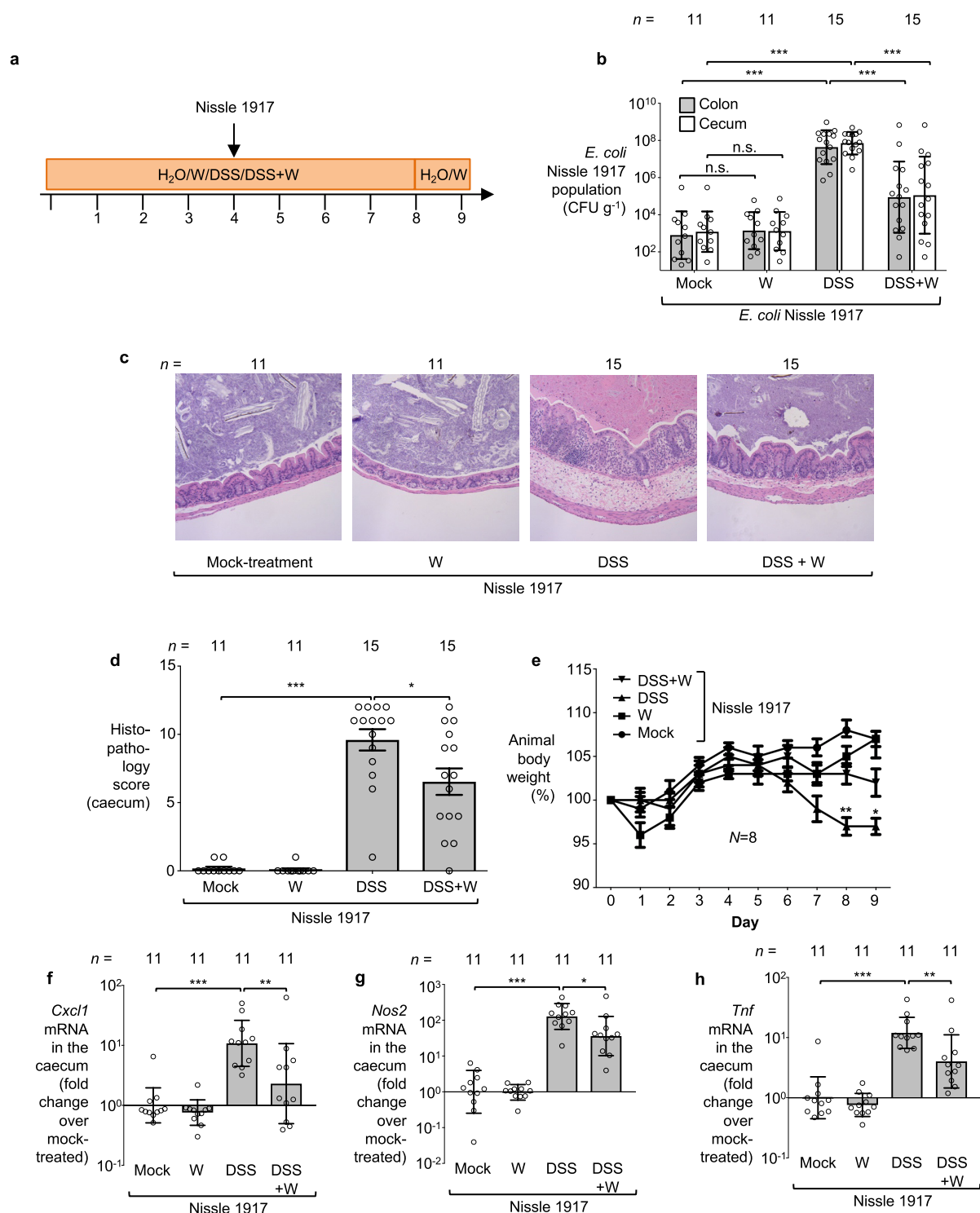
Extended Data Figure 1 | Effect of tungstate on anaerobic respiration *in vitro*. **a**, Nitrate reductase activity of *E. coli* Nissle 1917 measured in medium supplemented with sodium nitrate and the indicated concentrations of sodium tungstate. **b**, Competitive growth of the *E. coli* Nissle 1917 wild-type strain and the isogenic molybdenum-cofactor-deficient mutant in the presence of the indicated electron acceptors under

anaerobic conditions. **c**, Competitive growth of *E. coli* NRG857c wild-type strain and the isogenic molybdenum-cofactor-deficient mutant in the presence of the indicated electron acceptors under anaerobic conditions. In **a–c**, $n = 3$ replicates per condition; n denotes the number of biological replicates. Data are shown as geometric mean and geometric s.d. of three experiments.



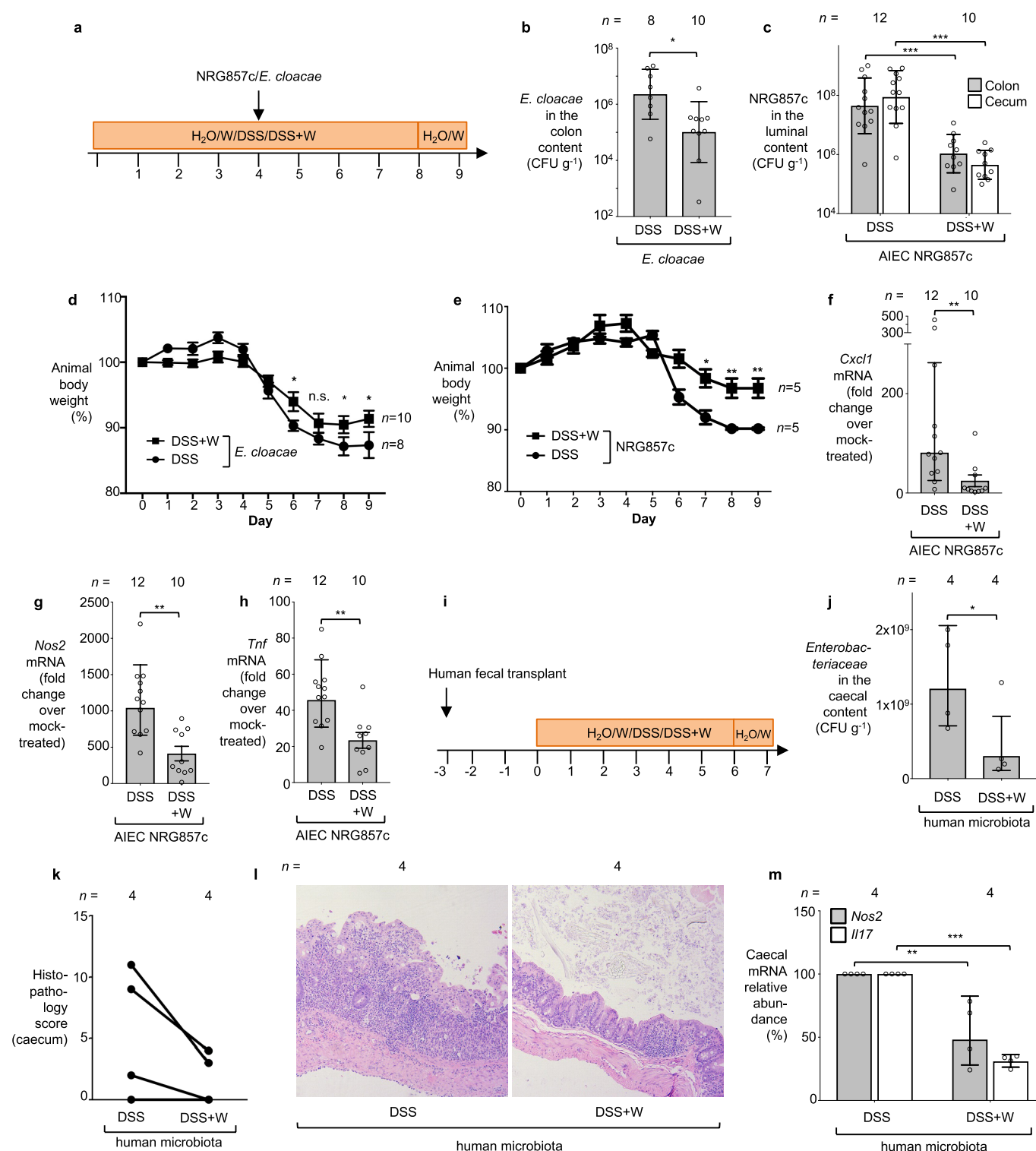
Extended Data Figure 2 | Tungstate treatment of mice colonized with *E. coli* K-12. Conventionally raised C57BL/6 mice were treated with 0.2% sodium tungstate, DSS, DSS plus sodium tungstate or left untreated (mock). After four days, animals were inoculated orally with the *E. coli* K-12 wild-type strain and the isogenic *moaA* mutant. C57BL/6 mice with naive microbiota (endogenous Enterobacteriaceae only) or germ-free C57BL/6 mice were treated similarly but were not inoculated with *E. coli* indicator strains. Samples were analysed after a total of nine days of treatment. **a, b**, Schematic representations of the colitis models. **c–h**, Transcription of *Nos2* (**c**), *Tnf* (**d**), *Il6* (**e**), *Lcn2* (**f**), *Cxcl2* (**g**) and

Ifng (**h**) in the caecal mucosa was determined by RT-qPCR. *E. coli* K-12: *n* = 6 per group. Endogenous Enterobacteriaceae: mock, *n* = 14; W, *n* = 6; DSS, *n* = 19; DSS+W, *n* = 19. Germ-free: mock, *n* = 5; DSS, *n* = 4; DSS+W, *n* = 5. **i**, Cumulative histopathology score for the colon tissue; mock, *n* = 14; W, *n* = 6; DSS, *n* = 19; DSS+W, *n* = 19; data are shown as mean and s.e.m. and each dot represents one animal. **j**, Colon length, *n* = 8 per group. **k**, Body weight of mice harbouring endogenous Enterobacteriaceae, *n* = 8 per group. **l**, Body weight of mice experimentally colonized with *E. coli* K-12, *n* = 6 per group. Unless otherwise noted, data are shown as geometric mean and geometric s.d.



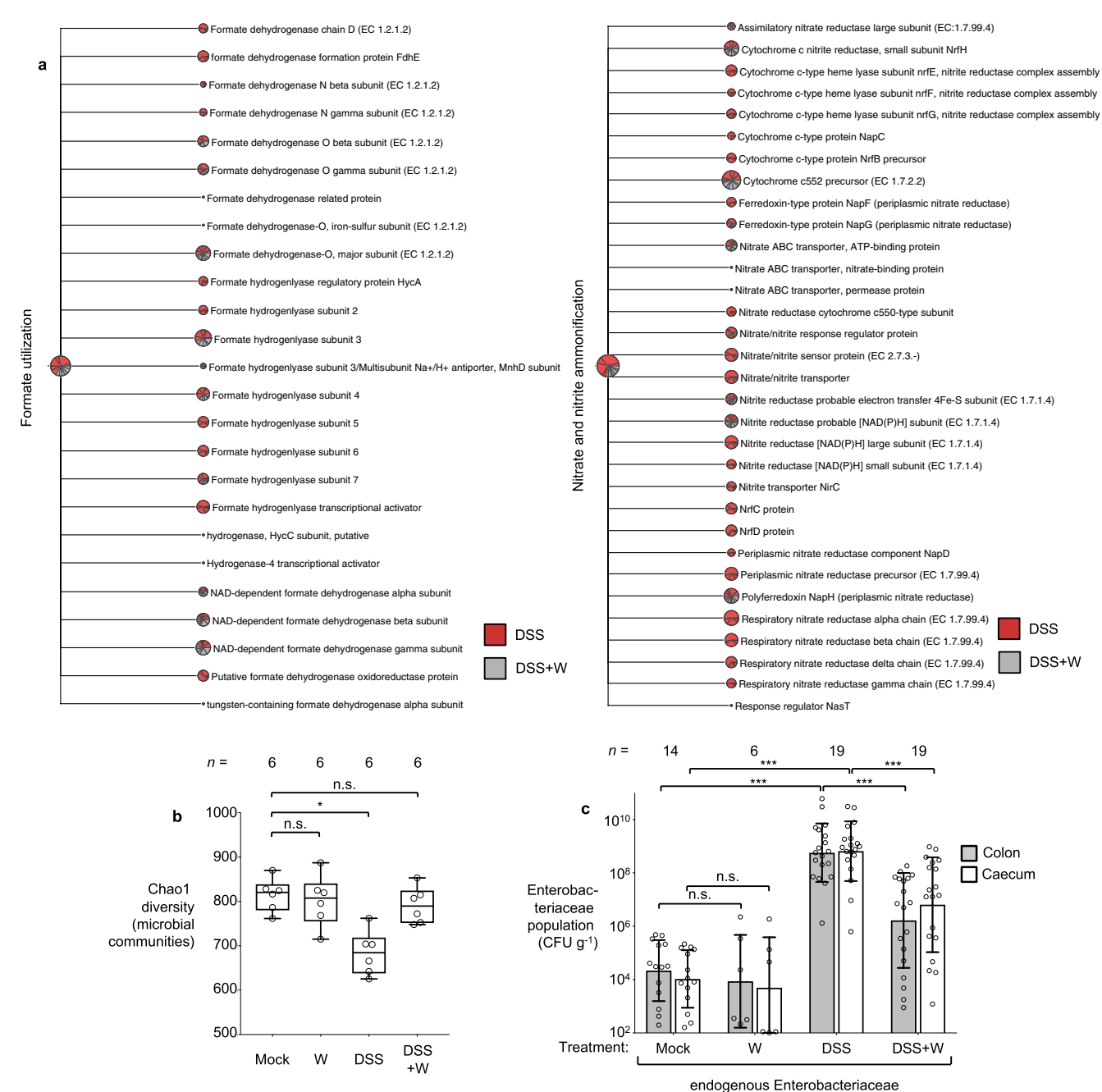
Extended Data Figure 3 | Effect of tungstate treatment on mice experimentally colonized with *E. coli* Nissle 1917. Groups of conventionally raised C57BL/6 mice were orally inoculated with the *E. coli* Nissle 1917 wild-type strain and treated with 0.2% sodium tungstate, DSS, DSS and sodium tungstate, or left untreated (mock) for nine days. **a**, Schematic representation of colitis model used in this figure. **b**, Bacterial load in the caecum (white bars) and colon content (grey bars). **c**, Formalin-fixed, haematoxylin and eosin-stained sections of the caecum were scored for the presence of inflammatory lesions; representative

images of stained caecal sections. **d**, Cumulative histopathology score for the caecum tissue; data are shown as mean and s.e.m., and each dot represents one animal. In **b–d**, mock and tungsten, $n = 11$ per group; DSS and DSS+W, $n = 15$ per group. **e**, Animal body weight, $n = 8$ per group. **f–h**, Transcription of the inflammatory marker genes *Cxcl1* (**f**), *Nos2*, (**g**) and *Tnf* (**h**) in the caecal mucosa was determined by RT-qPCR, $n = 11$ per group. Unless otherwise noted, data are shown as geometric mean and geometric s.d.



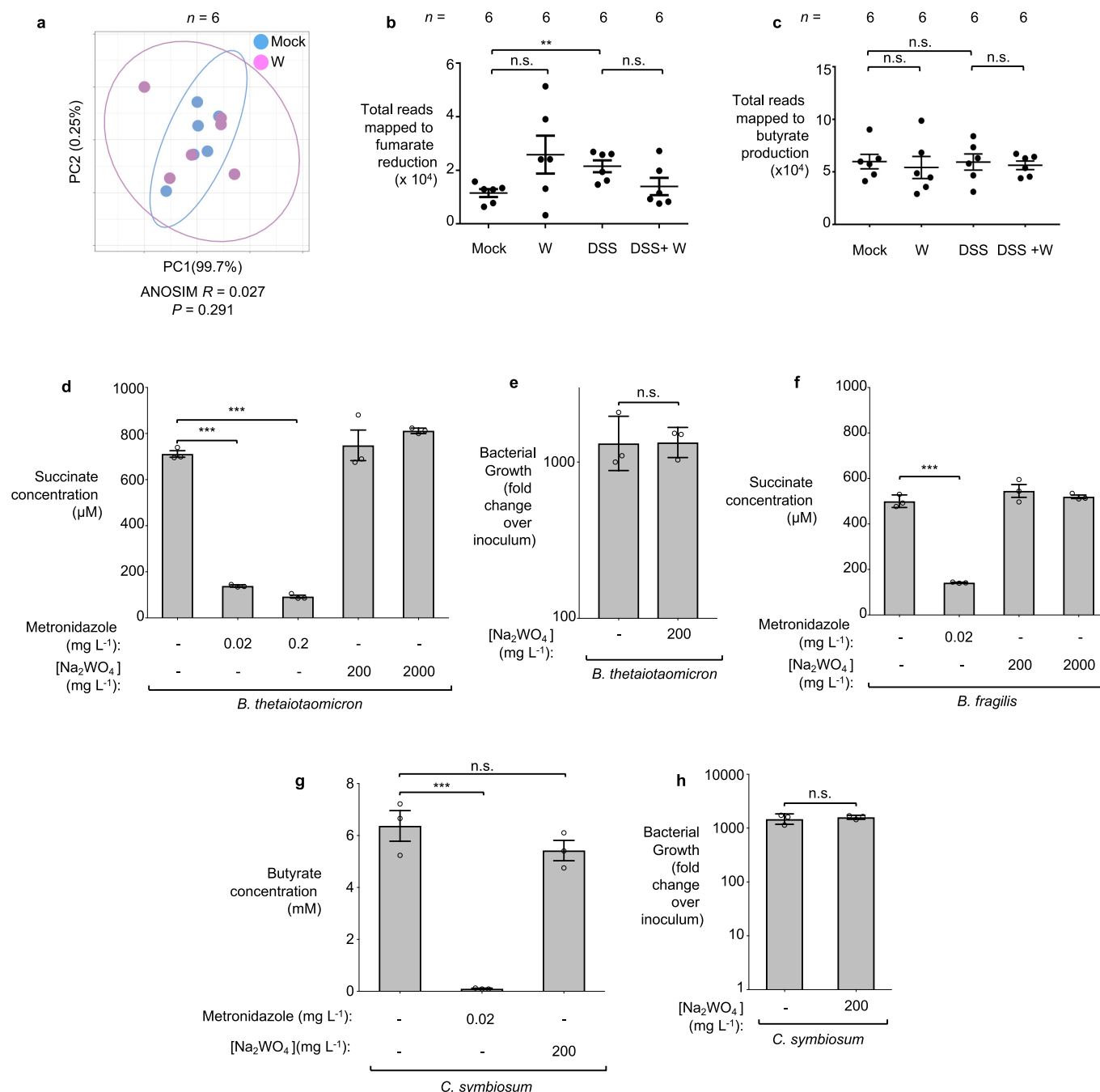
Extended Data Figure 4 | Effect of tungstate treatment on mice experimentally colonized with *Enterobacter cloacae* and adherent invasive *E. coli*. **a–h**, Conventionally raised C57BL/6 mice were treated with DSS or DSS plus tungstate. After four days, animals were inoculated intragastrically with the indicated bacterial strains. Samples were collected five days after inoculation. **a**, Schematic representation of the experiments. **b, c**, The total population of *E. cloacae* (**b**) and NRG857c (**c**; DSS, *n* = 12; DSS+W, *n* = 10) in the large intestinal content was determined by plating on selective medium. **d, e**, Animal body weight. In **b** and **d**: DSS, *n* = 8; DSS+W, *n* = 10. In **e**, *n* = 5 per group. **f–h**, Transcription levels of the inflammatory marker genes *Cxcl1* (**f**), *Nos2* (**g**) and *Tnf* (**h**) in the caecal

mucosa were determined by RT-qPCR; DSS, *n* = 12; DSS+W, *n* = 10. **i–m**, Paired germ-free Swiss-Webster mice received human faecal transplants and were treated with DSS or DSS plus 0.2% sodium tungstate for seven days; DSS, *n* = 4; DSS+W, *n* = 4. **i**, Schematic representation of the experiments. **j**, The abundance of *Enterobacteriaceae* in the caecal content was determined by plating on selective medium (MacConkey agar). **k**, Formalin-fixed, haematoxylin and eosin-stained sections of the mouse caecum were scored for the presence of inflammatory lesions; **l**, representative images. **m**, Transcription levels of the inflammatory marker genes *Nos2* and *Il17* in the mouse caecal mucosa. Data are shown as geometric mean and geometric s.d.



Extended Data Figure 5 | Effect of tungstate on the naive gut microbiome. Groups of C57BL/6 mice naturally harbouring Enterobacteriaceae were treated with 0.2% sodium tungstate, DSS, DSS plus tungstate or mock treatment for nine days (see also Extended Data Fig. 3b). **a**, Relative abundance of genes involved in formate and nitrate utilization in the caecal content shown by shotgun metagenomic sequencing (MEGAN5). Each section of the pie chart is representative of the number of mapped reads obtained for the individual animals ($n = 6$

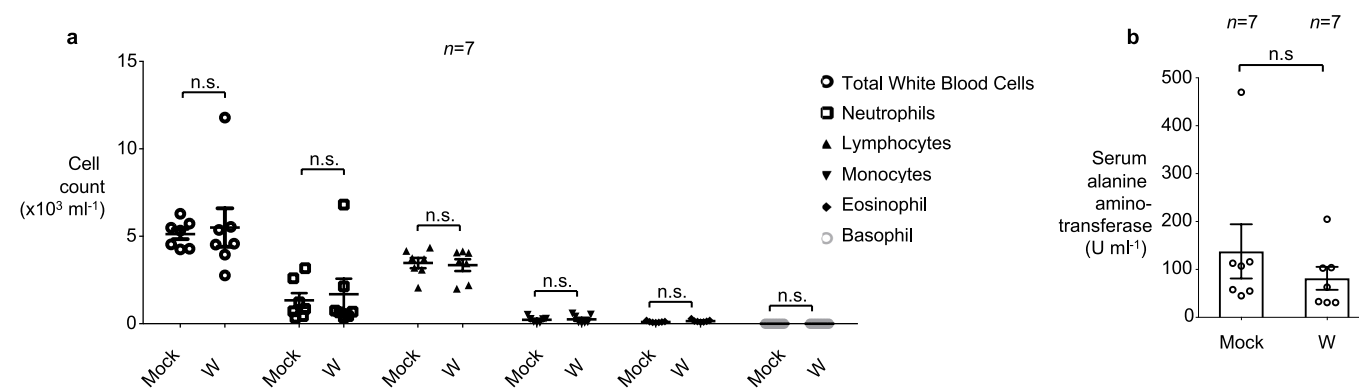
per group). **b**, Box-and-whisker plot (boxes show median, first and third quantiles, whisker denotes minimum to maximum range) of Chao1 alpha diversity of the caecal microbiota community based on 16S profiling ($n = 6$ per group). **c**, Abundance of endogenous Enterobacteriaceae family members determined by plating on selective medium (MacConkey agar, **c**): mock, $n = 14$; W, $n = 6$; DSS, $n = 19$; DSS+W, $n = 19$. Data are shown as geometric mean and geometric s.d.



Extended Data Figure 6 | Effect of tungstate treatment on obligate anaerobic commensal bacteria. **a–c**, Metagenomic analysis of the caecal content of mice described in Extended Data Fig. 3b. Principal coordinates analysis of global metabolic pathway (**a**) and quantification of reads involved in fumarate respiration (**b**) and butyrate production (**c**). Ellipses in **a** denote 95% confidence interval. Data are shown as mean and s.d.; $n = 6$ per group. **d–f**, *Bacteroides thetaiotaomicron* or *Bacteroides fragilis* were cultured anaerobically in mucin broth at 37 °C for 48 h. The medium was supplemented with sodium tungstate or metronidazole as indicated. Succinate production by *B. thetaiotaomicron* (**d**) and *B. fragilis* (**f**) was

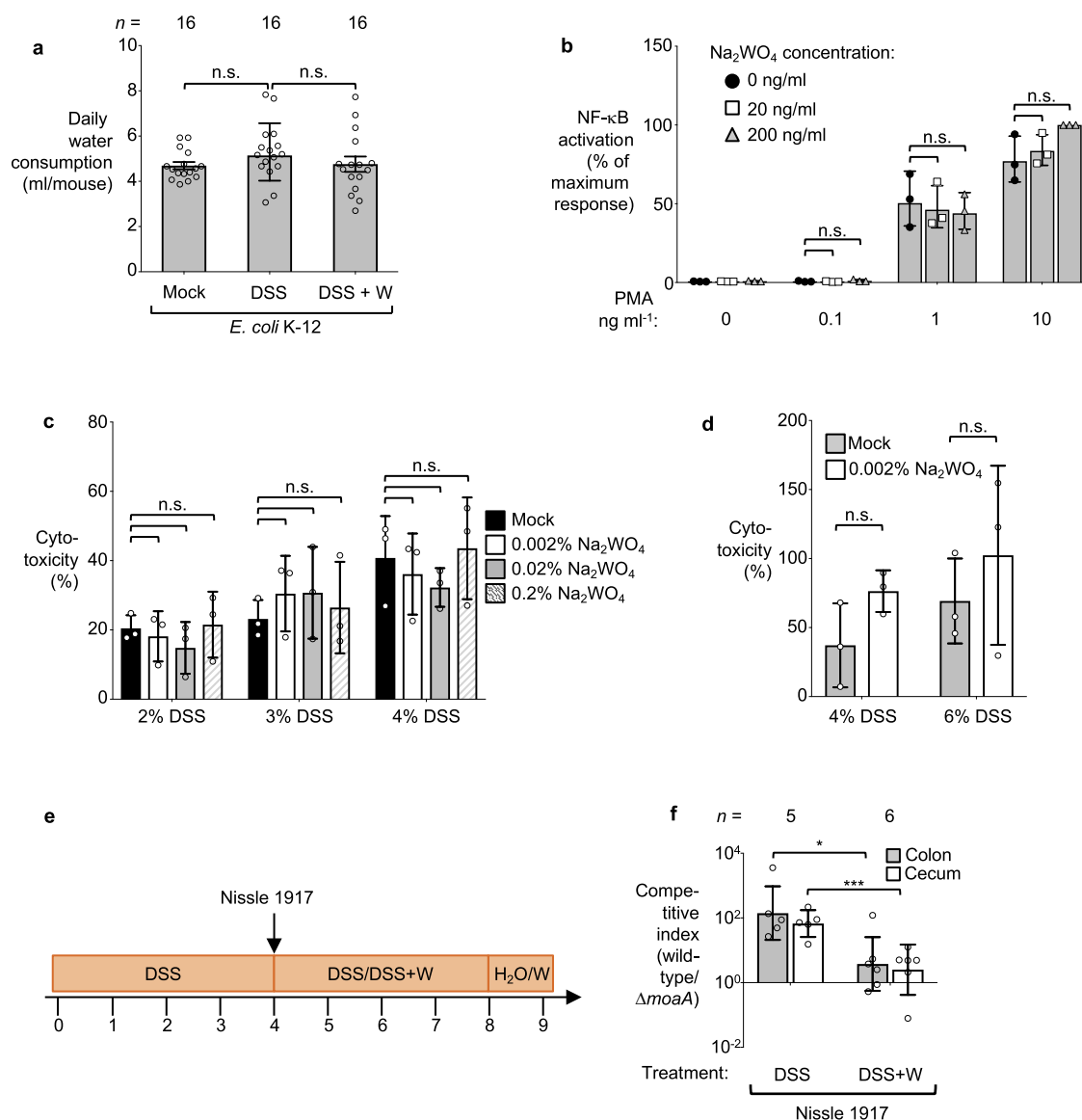
assessed by GC–MS. The growth of *B. thetaiotaomicron* was quantified by plating serial dilutions of bacterial culture on blood agar (**e**).

g, *C. symbiosum* was inoculated into chopped meat broth and incubated anaerobically at 37 °C for 36 h. Butyrate concentration in the medium was measured using GC–MS. **h**, *C. symbiosum* was cultured anaerobically in chopped meat broth at 37 °C for 48 h. The growth of *C. symbiosum* was determined by plating serial dilutions of bacterial culture on thioglycolate plates. $n = 3$ biological replicates per condition. Data are shown as geometric mean and geometric s.d. of three experiments.



Extended Data Figure 7 | Assessment of overall health of mice treated orally with tungstate. Groups of seven mice were either mock-treated or treated with 0.2% sodium tungstate in drinking water for nine days.

a, Complete blood count. **b**, Serum alanine amino-transferase concentration. $n = 7$ animals per group. Data are shown as geometric mean and geometric s.d.



Extended Data Figure 8 | Exposure of cultured cells to sodium tungstate. **a**, Daily water consumption of mice inoculated with *E. coli* K-12. Each dot represents the average daily water consumption (ml per day) of three mice, obtained at eight time points, with two cages per treatment group, $n = 16$. **b**, HeLa57A cells, expressing luciferase under the control of a NF- κ B-dependent promoter, were treated with PMA and sodium tungstate at the indicated concentrations. Relative luciferase activity was determined after 5 h. **c**, **d**, MODE-K or BMDMs cells were treated with DSS or DSS plus sodium tungstate at the indicated concentrations for 24 h. The release of lactate dehydrogenases into the culture supernatant by MODE-K cells (**c**) or BMDMs (**d**) was

measured. In **b–d**, $n = 3$ biological replicates per condition. **e**, **f**, Groups of conventionally raised C57BL/6 mice were treated with DSS for four days. Animals were inoculated intragastrically with an equal mixture of the indicated *E. coli* Nissle 1917 wild-type strain and an isogenic *moaA* mutant. On the day of inoculation, a subset of mice was switched to DSS plus sodium tungstate for four days while a control group remained on DSS treatment. Schematic representation of experiment (**e**). The competitive index in the caecal (white bars) and colon content (grey bars) was analysed 5 days after inoculation (**f**; DSS, $n = 5$; DSS+W, $n = 6$). Data are shown as geometric mean and geometric s.d.

Clonal analysis of lineage fate in native haematopoiesis

Alejo E. Rodriguez-Fraticelli^{1,2}, Samuel L. Wolock³, Caleb S. Weinreb³, Riccardo Panero⁴, Sachin H. Patel¹, Maja Jankovic¹, Jianlong Sun^{1,†}, Raffaele A. Calogero⁴, Allon M. Klein³ & Fernando D. Camargo^{1,2}

Haematopoiesis, the process of mature blood and immune cell production, is functionally organized as a hierarchy, with self-renewing haematopoietic stem cells and multipotent progenitor cells sitting at the very top^{1,2}. Multiple models have been proposed as to what the earliest lineage choices are in these primitive haematopoietic compartments, the cellular intermediates, and the resulting lineage trees that emerge from them^{3–10}. Given that the bulk of studies addressing lineage outcomes have been performed in the context of haematopoietic transplantation, current models of lineage branching are more likely to represent roadmaps of lineage potential than native fate. Here we use transposon tagging to clonally trace the fates of progenitors and stem cells in unperturbed haematopoiesis. Our results describe a distinct clonal roadmap in which the megakaryocyte lineage arises largely independently of other haematopoietic fates. Our data, combined with single-cell RNA sequencing, identify a functional hierarchy of unilineage- and oligolineage-producing clones within the multipotent progenitor population. Finally, our results demonstrate that traditionally defined long-term haematopoietic stem cells are a significant source of megakaryocyte-restricted progenitors, suggesting that the megakaryocyte lineage is the predominant native fate of long-term haematopoietic stem cells. Our study provides evidence for a substantially revised roadmap for unperturbed haematopoiesis, and highlights unique properties of multipotent progenitors and haematopoietic stem cells *in situ*.

To probe native lineage relationships in the fully unperturbed bone marrow, we used the Sleeping Beauty lineage-tracing model and TARIS (T7-amplification mediated recovery of integration sites), an improved transposon integration sequencing technique (Fig. 1a and Extended Data Figs 1 and 2)¹¹. Our analysis relies on comparing tags across multiple differentiated populations at different time points to understand the dynamics of lineage coupling, without the need to isolate and transplant prospective progenitor subsets (Fig. 1b). We pulsed adult Sleeping Beauty mice with doxycycline (Dox) for two days and, at one, two, four, and eight weeks after induction, sorted transposon-labelled (DsRed⁺) nucleated erythroblasts, megakaryocyte progenitors (MkP), granulocytes, monocytes, and B-cell progenitors (Fig. 1c). Notably, control experiments demonstrated that negligible amounts of transposition occur one day after removal of Dox (Extended Data Fig. 3).

We observed that blood lineages were mostly segregated for the first two weeks, suggesting their replacement by unilineage progenitors during this period (Fig. 1d). Later, we began to detect a significant number of shared tags across lineages, revealing the activity of oligolineage progenitors (Fig. 1d and Extended Data Fig. 4). At four weeks, 40.5% ($\pm 8.4\%$) of all monocyte-detected tags (approximately 289 \pm 89 clones) were also found in the granulocyte compartment, confirming their well-established common origin (Fig. 1e)⁴. Unexpectedly, a similar proportion of erythroblast clones were also found shared with

granulocyte/monocyte (myeloid) tags (Fig. 1d, e), revealing a common progenitor for erythrocytes, granulocytes, and monocytes at this stage. Remarkably, we detected virtually no MkP clones that were shared exclusively with erythroblasts during the whole period of observation, which would have been predicted had a megakaryocyte–erythroid progenitor (MEP)-like cell existed (Fig. 1d, e and Extended Data Fig. 4b)^{12,13}. At eight weeks, our analysis revealed the activity of a set of multilineage clones (239 \pm 58), with lymphoid (B-cell progenitor), granulocyte/monocyte, and erythroid contribution, but still with no presence in MkP, indicating the existence of megakaryocyte-deficient lympho-erythromyeloid progenitors (Fig. 1d, e). We did observe a very small (9.7 ± 2.8), yet increasing, number of MkP tags shared with multiple lineages after eight weeks (Fig. 1e and Extended Data Fig. 4a, b), suggesting that clonal megakaryocyte-lineage production can also be associated with multilineage outcomes, although at lower frequencies. Spearman's rank correlation analyses of tag-read distribution between lineage pairs showed a progressive association of granulocyte/monocyte, erythro-myeloid, and lympho-myeloid progenitors, segregated from MkP (Fig. 1f, g). To address potential sampling and sensitivity limitations, we performed independent TARIS amplifications (Extended Data Fig. 5) and clone-specific PCRs (not-shown). Taken together, our results provide evidence for novel lineage couplings during unperturbed haematopoiesis, in which the megakaryocyte lineage is produced largely independently from the other haematopoietic lineages, and argue for the robust activity of erythromyeloid and lympho-erythromyeloid progenitor clones.

We next aimed to identify ancestral relationships by comparing the clonal repertoires of differentiated cells and previously defined progenitor populations. Classically, oligopotent progenitors reside in the common myeloid progenitor (CMP), granulocyte–monocyte progenitor (GMP), and MEP phenotypic gates (referred together as myeloid progenitors, or MyPs)⁴. Our data revealed largely unilineage outcomes for detected MyPs (89.0 \pm 0.8%), suggesting that these populations represent a collection of lineage-restricted progenitors, functionally validating predictions from single-cell expression profiling (Extended Data Fig. 6)^{14–16}. We next focused on the multipotent progenitors (MPPs), the cellular subset proposed to be upstream of MyPs. At 1 and 2 weeks, we observed a small number of 'active' MPP tags (overlapping with Lin⁺ tags), which aligned mostly with single lineages (1 week: 75.8 \pm 5.0%; 2 weeks: 66.3 \pm 6.1%), suggesting the existence of a small population of lineage-committed MPPs that rapidly produce differentiated progeny (Fig. 2a, b and Extended Data Fig. 7a). MPP output significantly increased at 4–8 weeks for all lineages (9.35 \pm 0.6% of all MPP tags at 8 weeks), consisting mostly of oligolineage erythromyeloid clones (79.2 \pm 5.3% of active MPP clones). A robust number of lympho-erythromyeloid MPP clones (12 \pm 2) were detected beginning at eight weeks (Fig. 2a), consonant with our analysis of Lin⁺ fractions (Fig. 1f). Although we also observed oligolineage

¹Stem Cell Program, Boston Children's Hospital, Boston, Massachusetts 02115, USA. ²Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts 02138, USA. ³Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁴Molecular Biotechnology Center, Department of Clinical and Biological Sciences, University of Torino, Torino 10126, Italy. [†]Present address: School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China.

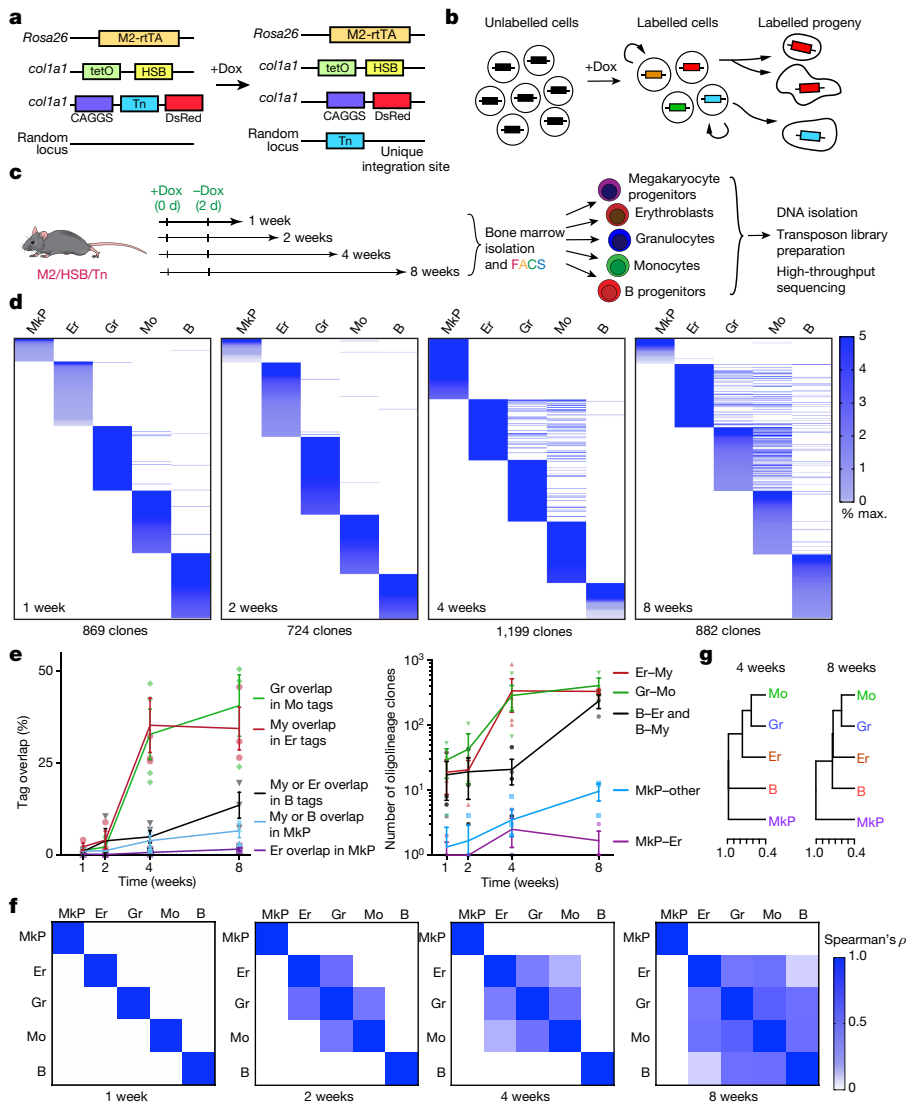


Figure 1 | Clonal analysis of haematopoietic lineage fates in the native bone marrow.

a, M2/HSB/Tn (transcriptional activator M2/hyperactive Sleeping Beauty/transposon) mouse model. Addition of Dox induces random transposition of the transposon, and concomitant cell labelling with DsRed. The transposon insertion site is stable after removal of Dox. **b**, Transposon lineage tracing. Shared tags can be detected between a self-renewing progenitor stem cell and its progeny, or between two different mature cell populations. **c**, Experimental design. M2/HSB/Tn mice were labelled with Dox for 2 days and five blood lineages were isolated from bone marrow after different periods of time. Transposon insertion tag libraries were prepared and sequenced for each population. **d**, Alignment of transposon tags from different lineage-committed (Lin^+) blood cell populations in the bone marrow at 1–8 weeks. Tags are coloured by frequency in each lineage, and organized by rank. Each chart is representative of three independent experiments. MkP, megakaryocyte progenitors; Er, erythroblasts; Gr, granulocytes; Mo, monocytes; B, B-cell progenitors. **e**, Percentage of clonal overlap between designated lineage pairs (left), and quantification of total number of detected bi-/tri-lineage clones at 1–8 weeks (right). Abbreviation My refers to either granulocyte or monocyte lineage. Mean \pm s.d. from three independent mice. **f**, Spearman's correlation coefficient (ρ) matrices for all Lin^+ tags at 1–8 weeks. Each matrix is the average from three independent experiments per time point. **g**, Hierarchical clustering of blood lineages using $(1 - \rho)$ as the distance measure (4 and 8 weeks after labelling).

MkP-producing MPP clones, MkP overlap was more lineage-restricted than any other lineage, even after eight weeks (MkP $67.8 \pm 8.0\%$ versus other $22.1 \pm 4.6\%$; Fig. 2a, b and Extended Data Fig. 7b), indicating that at least a subset of MPPs is responsible for a stable restricted contribution to the megakaryocyte lineage.

Our analyses also provided relative quantitative information about the dynamics of lineage replacement by MPPs. For instance, the average clone size of MPP-derived erythromyeloid clones at eight weeks was 18.3 ± 7.7 -fold larger than non-MPP-derived clones, suggesting a significant cellular amplification, in contrast to the B-cell progenitor lineage (1.2 ± 0.4 -fold; Fig. 2c). In addition, we found that the erythroid lineage was replaced at the fastest rate, with at least 35% of all erythroblast reads overlapping with MPPs after just two weeks, from just a handful of erythroid-committed MPPs (Fig. 2d, e). By comparison, the granulocyte/monocyte-producing MPPs achieved similar levels of replacement only after two months. Considering that our analysis cannot measure the contribution of MPP clones that disappear from the MPP pool (that is, by cell death or differentiation), our results probably underestimate the overall MPP contribution.

To provide further insight into the heterogeneity and hierarchy of the haematopoietic stem cell (HSC)/MPP compartment, we sorted subsets within these populations using previously described surface markers and interrogated their single-cell gene expression landscape using inDrop (Fig. 3a–c)^{9,17}. Louvain–Jaccard clustering analysis of transcriptomes resulted in 12 reproducibly distinct clusters (Fig. 3b). Most analysed cells (78.9% of all subsets combined) fitted into one of

three major clusters that we labelled as unprimed ('C1', 'C2', 'C3') on the basis of the lack of expression of lineage-restricted gene signatures (Supplementary Table 2 and Extended Data Figs 8 and 9). We also identified several primed clusters (21.1% of HSCs/MPPs) that formed branches defined by progressive expression of genes associated with lineage commitment (Fig. 3b–d, right). Predictably, cells indexed as long-term (LT)-HSCs and MPP1s (also known as short-term (ST)-HSCs) mostly fitted into the 'C1' (67.9%) and 'C2' (78.3%) clusters, respectively. By contrast, other MPP subsets displayed different degrees of heterogeneity. MPP2s contained the largest proportion of primed cells (59.3%), and MPP4s the least (13.2%) (Fig. 3c, d). MPP2s comprised a larger number of erythroid-primed (18.7%) and megakaryocyte-primed (21.9%) cells, whereas MPP3s contained a larger number of granulocyte/monocyte-primed cells (20.8%) (Fig. 3c, d and Extended Data Fig. 8b). Using transposon tracing, we confirmed that MPP2s presented a preference for MkP production, and generated less multilineage output ($5 \pm 5\%$ of all active clones) within the first week, where their immediate progeny is likely to be measured, compared with MPP3s and MPP4s ($40.17 \pm 11.4\%$) (Fig. 3e, f). Analysis of tags not arising from upstream progenitors at four weeks revealed similar findings (Fig. 3g, h). On the contrary, MPP4s produced most lympho-erythromyeloid and multilineage clones (Fig. 3h) and preferentially overlapped with MPP1/ST-HSCs, suggesting that at least a fraction of MPP4s represent direct activated progeny of MPP1/ST-HSCs (Fig. 3i). Combined, our data support the notion that a functional hierarchy, consisting of progenitors at varying degrees of lineage priming, exists already within HSCs/MPPs.

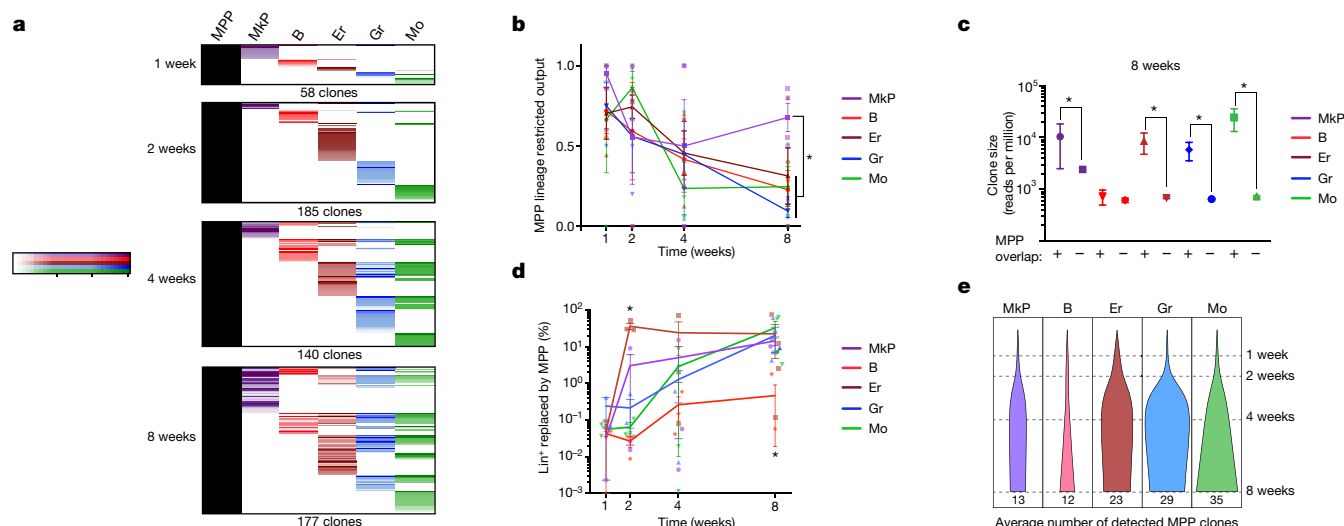


Figure 2 | Functional heterogeneity of MPP lineage fates in steady-state haematopoiesis. **a**, The alignment of all active MPP tags together with the five analysed blood lineages at each time point (all tags collected from three mice per time point). LT-HSC tags were analysed in parallel and excluded from the analysis to represent only MPP behaviour. **b**, Fraction of active MPP tags that overlap with a single lineage (calculated independently for each lineage). Values are mean \pm s.e.m. from three mice. * $P_{\text{MkP-Er}} = 0.13$, $P_{\text{MkP-Gr}} = 0.03$, $P_{\text{MkP-Mo}} = 0.03$, $P_{\text{MkP-B}} = 0.001$ (8 weeks). Abbreviations as Fig. 1. **c**, Distribution of Lin^+ clone sizes comparing tags overlapping with MPP versus non-overlapping at eight weeks. Values are

median and interquartile ranges of all detected clones from three mice. *Kolmogorov-Smirnov $P_{\text{MkP}} = 0.03$, $P_{\text{B}} = 0.25$, $P_{\text{Er}} = 0.03$, $P_{\text{Gr}} = 0.001$, $P_{\text{Mo}} = 0.003$. **d**, Fraction of each lineage replaced by MPPs calculated as the percentage of total MPP-overlapping lineage reads over time. Values are mean \pm s.e.m. from three independent mice. * $P_{\text{Er-Gr/Mo/B}} = 0.04$, $P_{\text{Er-MkP}} = 0.03$ (2 weeks) and $P_{\text{B-Er/MkP}} = 0.03$, $P_{\text{B-My}} = 0.04$ (8 weeks). **e**, Average number of detected active MPP clones per lineage per mouse at different time points (normalized for percentage DsRed labelling efficiency).

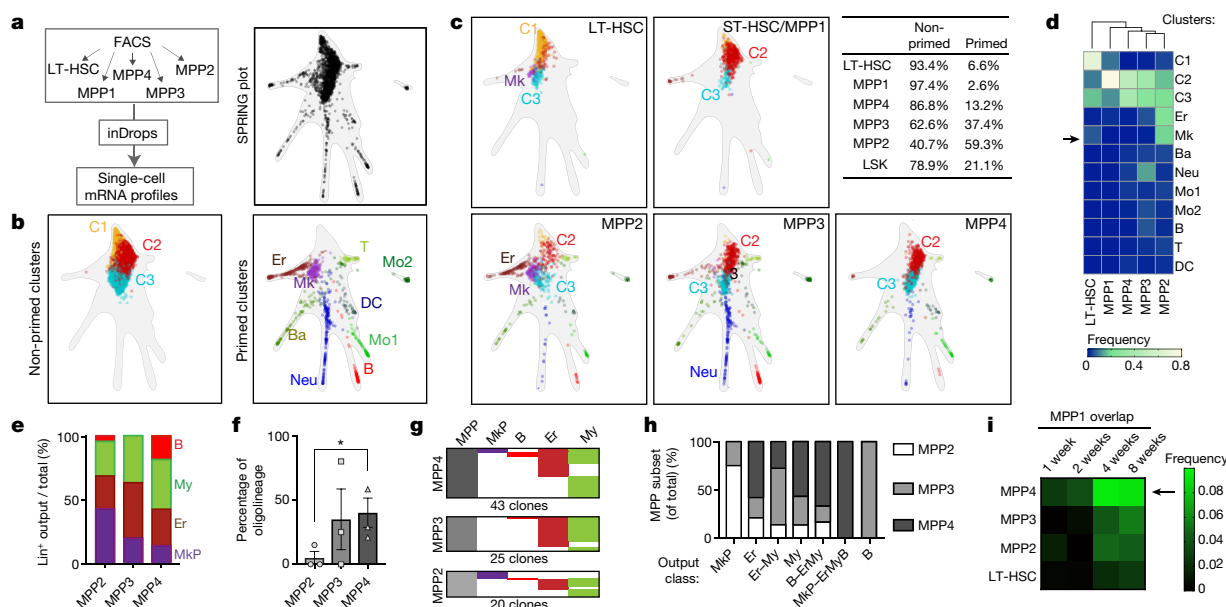


Figure 3 | Transcriptional and functional hierarchy of HSC and MPP subsets. **a**, Experimental design for inDrops experiment (left). Transcriptional fate map of combined fluorescence-activated cell sorting (FACS) subsets using the SPRING representation (subsampling *in silico* to represent proportions of the $\text{Lin}^- \text{Sca1}^+ \text{Kit}^+$ (LSK) gate). Points represent a single HSC/MPP distributed according to their similarity using gene expression variation. **b**, *In silico* identification of different cell populations within all combined HSC and MPP subsets. Non-primed clusters 1–3 (C1–C3, left) and lineage-primed clusters (right) are presented separated and labelled according to their primed lineage signatures: Neu, neutrophil; DC, dendritic cell; T, T-cell progenitor; B, B-cell progenitor; Er, erythroid; Mk, megakaryocyte; Mo1 and Mo2 represent two monocyte-like signatures. **c**, Plots showing localization of each sorted HSC/MPP subset within the combined SPRING plot. Top right, fraction of cells from each sorted HSC/MPP subtype (and LSK cells) that group within primed or non-primed clusters. **d**, Hierarchical clustering (Ward) of sorted

HSC/MPP subsets. For each FACS-sorted population, the fraction of cells corresponding to each cluster was used to analyse the similarity between subsets. The arrow points out the megakaryocyte-primed cluster within the LT-HSC gate. **e**, Fraction of lineage-restricted MPP-overlapping clones corresponding to each lineage, for each MPP subset at one week. Values are mean \pm s.e.m. of three independent mice. NS, not significant. **f**, Fraction of oligolineage output of each MPP subset after 1 week. Values are mean \pm s.e.m. of three independent mice. *Paired two-tailed *t*-test (MPP2 versus MPP4), $P = 0.033$. **g**, Alignment of Lin^+ progeny tags of different MPP subsets (excluding tags present in HSCs/MPP1s) at four weeks. **h**, Fraction corresponding to each MPP subset for each representative lineage fate (including restricted, oligolineage, and multilineage output) at four weeks (all tags detected from four mice). **i**, Frequency of MPP2/3/4 tags (and LT-HSC tags) overlapping with MPP1 at 1–8 weeks (average of three mice per time point).

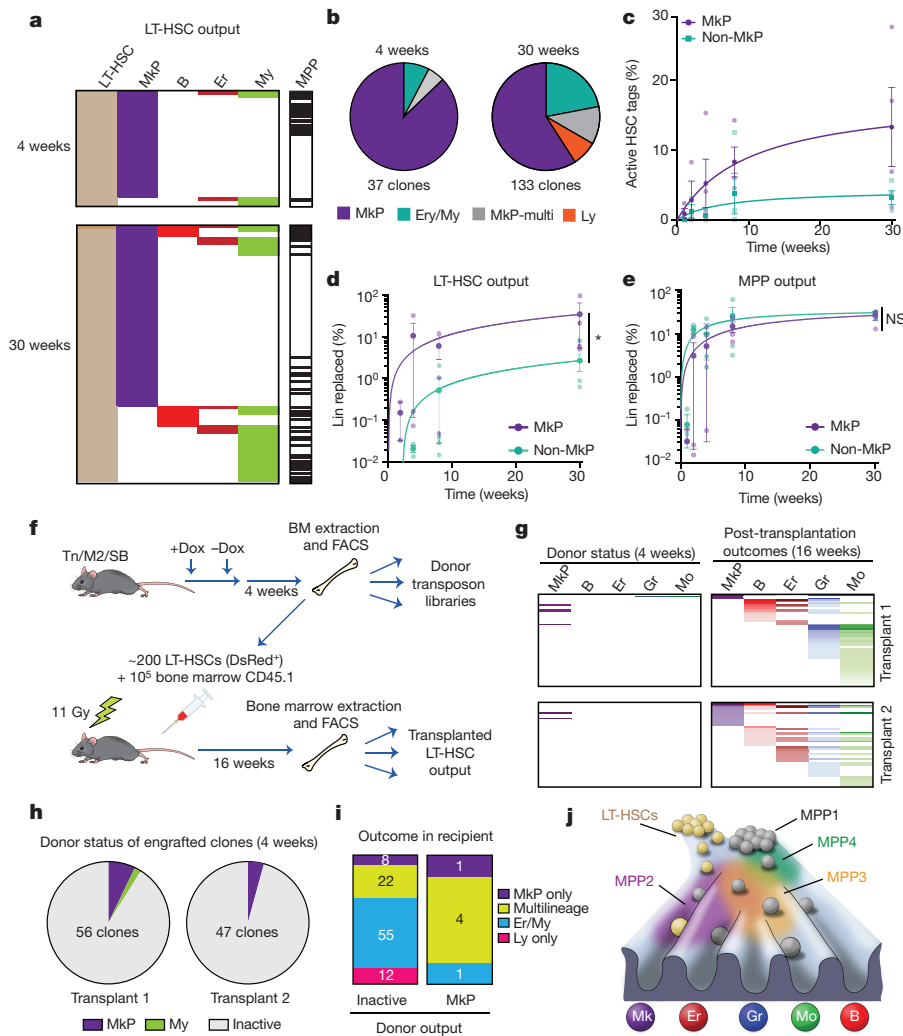


Figure 4 | Steady-state MkP output from bona fide LT-HSCs. **a**, LT-HSCs, MPPs, and Lin⁺ cells were purified from bone marrow at 4 and 30 weeks and their transposon tag content was analysed. Only the LT-HSC tags overlapping with detectable Lin⁺ progeny are shown. Abbreviations as Fig. 1. **b**, Distribution of types of progeny detected from LT-HSCs at 4 weeks and 30 weeks after labelling. Data are pooled from four independent M2/HSB/Tn mice per time point. Ly, lymphocyte. **c**, Percentage of labelled LT-HSC clones producing progeny at 1–8 weeks. Values are mean \pm s.e.m. of three or four independent mice. **d**, Dynamics of megakaryocyte versus non-megakaryocyte lineage replacement by LT-HSCs (measured as the percentage of overlapping/total Lin⁺ reads). Values are mean \pm s.e.m. of three or four independent mice. Ratio paired *t*-test, $P = 0.014$. **e**, Dynamics of megakaryocyte versus erythrocyte/granulocyte/monocyte lineage replacement by MPPs (measured as the percentage of overlapping/total Lin⁺ reads). Values are mean \pm s.e.m. of three or four

Our single-cell RNA sequencing data also revealed that a subset of marker-defined LT-HSCs exhibited megakaryocyte-lineage priming (Fig. 3c, d and Extended Data Fig. 9). This is in line with previous reports of multipotent, yet platelet-biased, subsets of LT-HSCs in the context of transplantation^{10,18–23}. However, the physiological relevance of this observation in native haematopoiesis is unknown. With these precedents, we analysed the Lin⁺ transposon tag overlap of sorted LT-HSCs. Although only a very small number of LT-HSC clones was active four weeks after labelling ($5.5 \pm 2.3\%$), remarkably a large majority of these clones were found exclusively in the MkP population (Fig. 4a, b and Extended Data Fig. 10a). This megakaryocyte-restricted output of LT-HSCs was more pronounced after 30 weeks post-labelling (MkP: $13.3 \pm 5.6\%$; lymphoid/erythroid/myeloid: $3.2 \pm 1.0\%$) (Fig. 4c). Quantitatively, LT-HSCs accounted for replacing at least 31% of the

independent mice. Ratio paired *t*-test, $P = 0.599$. **f**, Experimental design for parallel analysis of native versus transplant output of LT-HSC clones. **g**, Alignment of all post-transplantation LT-HSC-derived lineages with unperturbed donor lineage tags. **h**, Pie-chart distribution of successfully engrafted LT-HSC clones by donor behaviour. Only megakaryocyte-restricted and granulocyte/monocyte-restricted output was observed. Inactive means non-detectable output in the donor. **i**, Post-transplantation outcomes comparing donor-inactive versus MkP-producing LT-HSC clones. **j**, Lineage fate landscape of unperturbed haematopoiesis. Self-renewing LT-HSCs preferentially replace the megakaryocyte lineage under steady state, and principally contribute to other blood lineages during transplantation or after injury. By contrast, MPPs take care of most steady-state lymphocyte, erythrocyte, and granulocyte/monocyte blood production. Different MPP sorting gates enrich for heterogeneous collections of lineage-primed and unprimed cell states within a continuum of lineage commitment pathways.

total MkP pool, compared with just 3.8% of granulocyte/monocyte and erythroblast reads (Fig. 4d). Among all MkP that had a detectable tag in primitive populations, approximately half demonstrated overlap with LT-HSCs and the other half with MPPs (where no LT-HSC tag was detected) (Extended Data Fig. 10b). MPP-overlapping clones contributed to the megakaryocyte lineage to a similar extent as LT-HSCs, markedly differing from lympho-erythromyeloid output, which is predominantly MPP-driven (Fig. 4e and Extended Data Fig. 10c). Our analyses also revealed that many LT-HSCs contribute to MkP in the absence of any intermediates in the MPP compartment (Fig. 4a), suggesting that at least a subset of LT-HSCs generates megakaryocyte lineage cells through a ‘direct’ pathway.

Previous studies have shown that the commonly used LT-HSC gate contains unilineage CD41⁺ megakaryocyte-restricted progenitors as

assayed by transplant or culture^{10,22}. To rule out potential contamination by such cells, we aimed to determine whether megakaryocyte-producing LT-HSC clones *in situ* had properties of classic LT-HSCs in the context of transplantation. For this, we transplanted clonally labelled LT-HSCs isolated from mice four weeks after induction, and at 16 weeks post-transplantation we purified mature lineages from recipients and compared their transposon repertoires with those of cells initially isolated from the donor (Fig. 4f). We observed that six out of eight detected megakaryocyte lineage-restricted LT-HSC clones in the donor were able to generate multilineage progeny in recipients (Fig. 4g, i). We reached similar conclusions when evaluating the culture potential of *in situ* MkP-producing LT-HSC clones (Extended Data Fig. 10d, e). Additionally, our results demonstrate that MkP production is not exclusive to the CD41⁺ LT-HSC fraction (Extended Data Fig. 10f, g). Thus, we conclude that most megakaryocyte lineage-producing clones residing in the LT-HSC gate are not simply megakaryocyte-restricted progenitors, but clones that can exhibit multipotency upon transplantation.

Our work here uncovers critical features of the native haematopoietic process. In our model, as much as half of the megakaryocyte lineage is produced independently of other lineages by cells at the top of the haematopoietic ladder (Fig. 4j). A heterogeneous hierarchy of lineage-restricted and oligolineage progenitors, historically classified as MPPs, produce other haematopoietic lineages with selective lineage couplings. Although our work still supports a model for progressive restriction of developmental potential, it suggests that these events are clonally heterogeneous and occur much earlier in the haematopoietic hierarchy, in line with recent data^{7,8,14,16}. Although our data fail to provide any evidence for CMP or MEP fates *in situ*, many experiments have provided evidence for MEP-like cells at a clonal level^{4,12,13,24}. We posit that while megakaryocyte–erythrocyte bipotential exists in transplant or culture settings, this fate is not substantially manifested in unperturbed conditions. Alternatively, such cellular behaviour might be too transient to be captured with our technology.

Our data demonstrate that at least a fraction of LT-HSCs behave as a potent source of MkP, indicating that the megakaryocyte fate is the predominant fate of HSCs *in situ*. However, these same cells exhibit potential for multilineage outcomes following transplantation. Thus, our findings highlight the critical differences between studying native fate versus potential in stem cell biology. Although we are unable to conclude whether a particular subset or all LT-HSCs will eventually display megakaryocyte-producing behaviour, we favour the idea that most LT-HSC clones transition through a megakaryocyte-primed state with age. Our data also suggest that an MPP population (within MPP2) is involved in megakaryocyte production. It remains to be determined whether these represent two different pathways for megakaryocyte production or whether LT-HSCs are upstream of MPP2s. Finally, our results are still consonant with the idea that adult LT-HSCs have a limited lympho-erythromyeloid output during steady state^{11,25}, although this finding has been debated²⁶. Future work with second-generation cell barcoding strategies^{27,28} in combination with Cre-based labelling will be needed to elucidate full lineage histories and determine the mechanisms of fate restriction.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 18 November 2016; accepted 21 November 2017.

Published online 3 January 2018.

- Morrison, S. J., Wandycz, A. M., Hemmati, H. D., Wright, D. E. & Weissman, I. L. Identification of a lineage of multipotent hematopoietic progenitors. *Development* **124**, 1929–1939 (1997).
- Morrison, S. J. & Weissman, I. L. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity* **1**, 661–673 (1994).
- Adolfsson, J. *et al.* Identification of Flt3⁺ lympho-myeloid stem cells lacking erythro-megakaryocytic potential: a revised road map for adult blood lineage commitment. *Cell* **121**, 295–306 (2005).

- Akashi, K., Traver, D., Miyamoto, T. & Weissman, I. L. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404**, 193–197 (2000).
- Ceredig, R., Rolink, A. G. & Brown, G. Models of haematopoiesis: seeing the wood for the trees. *Nat. Rev. Immunol.* **9**, 293–300 (2009).
- Forsberg, E. C., Serwold, T., Kogan, S., Weissman, I. L. & Passegué, E. New evidence supporting megakaryocyte–erythrocyte potential of flk2/flt3⁺ multipotent hematopoietic progenitors. *Cell* **126**, 415–426 (2006).
- Notta, F. *et al.* Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* **351**, aab2116 (2016).
- Perié, L., Duffy, K. R., Kok, L., de Boer, R. J. & Schumacher, T. N. The branching point in erythro-myeloid differentiation. *Cell* **163**, 1655–1662 (2015).
- Pietras, E. M. *et al.* Functionally distinct subsets of lineage-biased multipotent progenitors control blood production in normal and regenerative conditions. *Cell Stem Cell* **17**, 35–46 (2015).
- Yamamoto, R. *et al.* Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* **154**, 1112–1126 (2013).
- Sun, J. *et al.* Clonal dynamics of native haematopoiesis. *Nature* **514**, 322–327 (2014).
- Debili, N. *et al.* Characterization of a bipotent erythro-megakaryocytic progenitor in human bone marrow. *Blood* **88**, 1284–1296 (1996).
- Sanada, C. *et al.* Adult human megakaryocyte–erythroid progenitors are in the CD34⁺CD38^{mid} fraction. *Blood* **128**, 923–933 (2016).
- Paul, F. *et al.* Transcriptional heterogeneity and lineage commitment in myeloid progenitors. *Cell* **163**, 1663–1677 (2015).
- Pronk, C. J. *et al.* Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell* **1**, 428–442 (2007).
- Velten, L. *et al.* Human haematopoietic stem cell lineage commitment is a continuous process. *Nat. Cell Biol.* **19**, 271–281 (2017).
- Klein, A. M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
- Calaminus, S. D. *et al.* Lineage tracing of Ptf4-Cre marks hematopoietic stem cells and their progeny. *PLoS ONE* **7**, e51361 (2012).
- Gekas, C. & Graf, T. CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. *Blood* **121**, 4463–4472 (2013).
- Haas, S. *et al.* Inflammation-induced emergency megakaryopoiesis driven by hematopoietic stem cell-like megakaryocyte progenitors. *Cell Stem Cell* **17**, 422–434 (2015).
- Nishikii, H. *et al.* Unipotent megakaryopoietic pathway bridging hematopoietic stem cells and mature megakaryocytes. *Stem Cells* **33**, 2196–2207 (2015).
- Roch, A., Trachsel, V. & Lutolf, M. P. Single-cell analysis reveals cell division-independent emergence of megakaryocytes from phenotypic hematopoietic stem cells. *Stem Cells* **33**, 3152–3157 (2015).
- Sanjuan-Pla, A. *et al.* Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* **502**, 232–236 (2013).
- Vannucchi, A. M. *et al.* Identification and characterization of a bipotent (erythroid and megakaryocytic) cell precursor from the spleen of phenylhydrazine-treated mice. *Blood* **95**, 2559–2568 (2000).
- Busch, K. *et al.* Fundamental properties of unperturbed haematopoiesis from stem cells *in vivo*. *Nature* **518**, 542–546 (2015).
- Sawai, C. M. *et al.* Hematopoietic stem cells are the major source of multilineage hematopoiesis in adult animals. *Immunity* **45**, 597–609 (2016).
- Junker, J. P. *et al.* Massively parallel clonal analysis using CRISPR/Cas9 induced genetic scars. Preprint at <https://doi.org/10.1101/056499> (2017).
- Raj, B. *et al.* Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain by scGESTALT. Preprint at <https://doi.org/10.1101/205534> (2017).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to members of the Camargo and Klein laboratory for comments. A.R.F. is a Merck Fellow of the Life Sciences Research Foundation and a non-stipendiary European Molecular Biology Organization postdoctoral fellow. This work was supported by National Institutes of Health grants HL128850-01A1 and P01HL13147 to F.D.C. F.D.C. is a Leukemia and Lymphoma Society and a Howard Hughes Medical Institute Scholar. A.M.K. is supported by a Burroughs-Wellcome Fund CASI award, and by the Edward J. Mallinckrodt Fellowship.

Author Contributions A.R.F. and F.D.C. designed the study, analysed the data, and wrote the manuscript. A.R.F. performed and analysed the experiments, assisted by M.J., S.P., and J.S. S.W., C.W., R.P., R.A.C., and A.M.K. designed and analysed inDrops experiments and transcriptome data. F.D.C. supervised the study.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to F.D.C. (Fernando.Camargo@childrens.harvard.edu).

Reviewer Information Nature thanks B. Gottgens and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Mice. The M2/HSB/Tn mice were generated as previously described¹¹. To induce transposon mobilization, 8- to 10-week-old male or female mice with the M2/HSB/Tn genotype were fed with 2 mg ml⁻¹ Dox together with 5 mg ml⁻¹ sucrose in drinking water for 48 h. Thereafter, Dox was removed and successful labelling was verified by retro-orbital sinus peripheral blood collection and analysis (70 µl) after 1 week. All animal procedures were approved by the Boston Children's Hospital Institutional Animal Care and Use Committee. Previous studies have estimated that most haematopoietic lineages are replaced by MPPs within 1–2 months after label^{25,29–31}. Thus, for Lin⁺ lineage coupling studies, M2/HSB/Tn mice were analysed within the first 8 weeks after labelling. Since MyPs have limited self-renewal capacity and are rapidly replaced by MPPs, we performed the MyP analysis at short time points after labelling (1 week) and only considered transposon tags not simultaneously present in MPPs.

Bone marrow preparation. After euthanasia, whole bone marrow (excluding the cranium) was immediately isolated in 2% fetal bovine serum in phosphate buffered saline, and erythrocytes were removed with red blood cell lysis buffer. CD45.1 (Ly5.1) mice were used as transplantation recipients (B6.SJL-Ptprca Pep3b/BoyJ, stock 002014, the Jackson Laboratory).

FACS. Lineage depletion was performed using magnetic-assisted cell sorting (Miltenyi Biotec) with anti-biotin magnetic beads and the following biotin-conjugated lineage markers: CD3e, CD19, Gr1, Mac1, and Ter119. Cell populations from bone marrow were purified through four-way sorting using FACSaria (Becton Dickinson) and six-way sorting using MoFlo XDP (Beckman Coulter). The following combinations of cell surface markers were used to define these cell populations. Erythroblasts: 7/4⁺Ly6G⁺Ter119⁺CD71⁺FSC^{hi}; granulocytes: Ly6G⁺7-4⁺B220⁻Ter119⁻; monocytes: Ly6G⁻7/4⁺B220⁻Ter119⁻; pro-/pre-B cells: Ly6G⁻B220⁺IL7Ra⁺; MkP: Lin⁻Kit⁺Sca1⁻CD150⁺CD41⁺; MPP1/ST-HSC: Lin⁻Kit⁺Sca1⁺Flt3⁻CD150⁻CD48⁻; MPP2: Lin⁻Kit⁺Sca1⁺Flt3⁻CD150⁺CD48⁺; MPP3: Lin⁻Kit⁺Sca1⁺Flt3⁻CD150⁻CD48⁺; MPP4: Lin⁻Kit⁺Sca1⁺Flt3⁺CD48⁺; LT-HSC: Lin⁻Kit⁺Sca1⁺Flt3⁻CD150⁺CD48⁻ (\pm CD41). Other populations are defined in Supplementary Table 1. Representative examples of sorted populations are shown in Supplementary Figs 1–3. Flow cytometry data were analysed with FlowJo (Tree Star). For transposon tag content extraction and analysis, we FACS-sorted all the available cells from the whole bone marrow extract using purity modes (approximately 98% purity) at about 75–80% efficiency. The antibodies (their clone number, the commercial house, and concentration) were as follows: Ly6B.2 FITC (7/4, Miltenyi, 1:100), Ly6G Alexa Fluor 700 (1A8, eBiosciences, 1:50), Ter119 APC (TER119, eBiosciences, 1:100), CD71 BV510 (C2, BD biosciences, 1:100), CD45R(B220) eFluor 450 (RA3-6B2, eBiosciences, 1:100), CD19 APC/Cy7 (1D3, eBiosciences, 1:50), CD127(IL-7R α) PE/Cy7 (A7R34, Biologend, 1:25), CD117 (Kit) FITC/APC (2B8, eBiosciences, 1:100), Ly6a (Sca1) PE/Cy7 (D7, eBiosciences, 1:100), CD135 (Flt3) APC (A2F10, Biologend, 1:25), CD150 PE/Cy5 (TC15-12F12.2, Biologend, 1:100), CD48 APC/Cy7 (HM48-1, BD biosciences, 1:100), CD41 BV605 (MwReg30, Biologend, 1:100), CD3e biotin (145-2C11, eBiosciences, 1:100), CD19 biotin (MB19-1, eBiosciences, 1:100), Gr1 biotin (RB6-685, eBiosciences, 1:100), CD11b (Mac1) biotin (M1/70, eBiosciences, 1:100), Ter119 biotin (TER119, eBiosciences, 1:100), streptavidin eFluor 450 (eBiosciences, 1:200), FcgrII/III eFluor 450 (93, eBiosciences, 1:100), CD34-FITC (RAM, eBiosciences, 1:25), CD42 APC (HIP1, Biologend, 1:100), CD9 PE (MZ3, Biologend, 1:200).

Transplantation assays. Whole bone marrow cells or sort-purified LT-HSCs from M2/HSB/Tn mice were transplanted in 150 μ l of α MEM (Gibco, Thermo Fisher Scientific) through retro-orbital injection into γ -irradiated recipient mice (split dose of 2.5 + 2.5 Gy for sublethal irradiation, and 5.5 + 5.5 Gy for lethal irradiation, with a 2 h interval). Donor cell engraftment and label frequency were analysed after 16 weeks using LSRII equipment (Becton Dickinson).

HSC culture assays. One thousand sort-purified LT-HSCs from M2/HSB/Tn mice were cultured together with 10,000 MS-5 stromal cells in round-bottom 96-well plates together with SCF (100 ng ml⁻¹), TPO (100 ng ml⁻¹), Flt3L (50 ng ml⁻¹), IL7 (20 ng ml⁻¹), IL3 (10 ng ml⁻¹), IL11 (50 ng ml⁻¹), and GM-CSF (20 ng ml⁻¹) in α MEM with 1% penicillin/streptomycin and 10% FCS (Thermo Fisher) for 2 weeks, changing the medium 24 h after sort and then every 48 h (Becton Dickinson). Myeloid and lymphoid HSC progeny was FACS-sorted after labelling with Gr-1, Mac-1, CD19, and B220 antibodies (eBiosciences). All growth factors and cytokines were mouse recombinant and purchased from Peprotech.

DNA isolation and amplification. Cells of interest were sorted into 1.7-ml tubes and concentrated into 5–10 μ l of buffer by low-speed centrifugation (700g for 5 min). Samples with fewer than 10,000 cells were subjected to whole-genome amplification with a Phi29 kit (Epicentre/Lucigen) according to the manufacturer's

instructions. Samples with more than 10,000 cells were purified by a QIAamp DNA Micro kit (56304, Qiagen).

TARIS. Our original technique for molecular identification of transposon integration sites was based on ligation-mediated (LM) PCRs. Others and we have observed significant tag amplification biases with this method, which limit the quantitative potential of the clonal data obtained^{11,32,33}. To improve the current technique, we developed a method based on TARIS (Extended Data Fig. 1). This method provided similar sensitivity levels as LM-PCR but more quantitatively and reproducibly captures the clonal composition of complex samples (Extended Data Fig. 2). For TARIS, the total purified DNA was subjected to enzymatic restriction with 10 U of HindIII-HF (NEB) overnight. TARIS adaptor primer was hybridized and extended using 1 U Klenow DNA polymerase (NEB) for 2 h. Then, total DNA was cleaned up using AMPure XP SPRI beads (Beckman Coulter) and used as a template for a 20 μ l T7 RNA polymerization reaction (NEB, High Yield Hicribes T7 kit) overnight. Then, the template was digested with 1 U of Turbo DNase (Ambion) and the RNA product was polyadenylated using 1 U of poly A RNA polymerase (NEB). The poly A RNA was purified with SPRI beads, and then converted into cDNA using iScript reverse transcriptase (Biorad). TARIS cDNA was used as template for 30 PCR cycles using the HSB-transposon-specific Tn-1C, the MAF-Tn-1F, and the MAR-polyT primers for 30 cycles, and then 12 cycles of indexing PCR using the MP1 and ID primers (ID1-48) and the KAPA HiFi PCR kit. Solexa sequencing was performed on HiSeq 2000 (Illumina) at the Tufts Genomics Core. Tag identification and alignment was performed as previously described¹¹. In brief, we extracted the transposon-containing reads from each fastq file, trimmed the adaptor and transposon sequences, and aligned the integration sites to the reference mouse genome (Ensembl mm9) using Bowtie 1.2. Then, reads were normalized between samples (per million reads). Sequences were always compared with at least one additional independently labelled mouse, with libraries prepared in parallel and sequenced in the same HiSeq lane to account for contaminations. Tags present in the control mouse samples were filtered out (contaminating reads). Next, read frequencies were column-normalized, and graphs were coloured using a logarithmic scale. For hierarchical clustering based on transposon tag distribution, we first determined the Spearman's correlation matrix for the compared populations and then performed agglomerative clustering (single method) using (1 - correlation coefficient) as the distance metric. Curve fitting was performed with the Lowess function. All indicated statistical tests were two-tailed parametric *t*-tests using Welch's s.d. correction (exceptions are mentioned where appropriate). Data visualization and statistical analysis was performed using Microsoft Excel, R (version 3.3.1), and GraphPad Prism (version 7). Primers used were TARIS adaptor primer (5'-GCATTAGCGGCCGCAAATTAATACGACTCACTATAGGAGTCTAAAGCCATGACATC-3'), Tn1-C primer (5'-CTTGTGTCATGCACAAAGTAGATGTCC-3'), MAF-Tn1-1F primer (5'-ACACTCTTTCCCTACACGACGTCTTCCGATCTNNNCGAGTTTAAATGACTCCAAC-3'), and MAR-polyT primer (5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT-3'). All primers were ordered from IDT DNA technologies, at 100 nmole scale and HPLC-purified.

Single-cell RNA sequencing and low-level data processing. Transcriptome barcoding and preparation of libraries for single-cell mRNA sequencing was performed using the most up-to-date inDrops protocol³⁴. For our experiment the Lin-Sca1⁺Kit⁺ bone marrow fraction from a single BL6 mouse was labelled and FACS-sorted to purify the entire LT-HSC, MPP1, MPP2, MPP3, and MPP4 fractions. Approximately 2,000 cells of each fraction were encapsulated and libraries for all the populations were prepared the same day, with the same stock of primer-gels and RT-mix. Libraries were sequenced on an Illumina NextSeq 500 sequencer using a NextSeq High 75 cycle kit: 35 cycles for read 1, 6 cycles for index i7 read, and 51 cycles for read 2. Raw sequencing reads were processed using the inDrop pipeline previously described, with the following modifications: Bowtie version 1.1.1 was used with parameter $-e\ 100$; all ambiguously mapped reads were excluded from analysis; and reads were aligned to the Ensembl release 81 mouse mm10 cDNA reference.

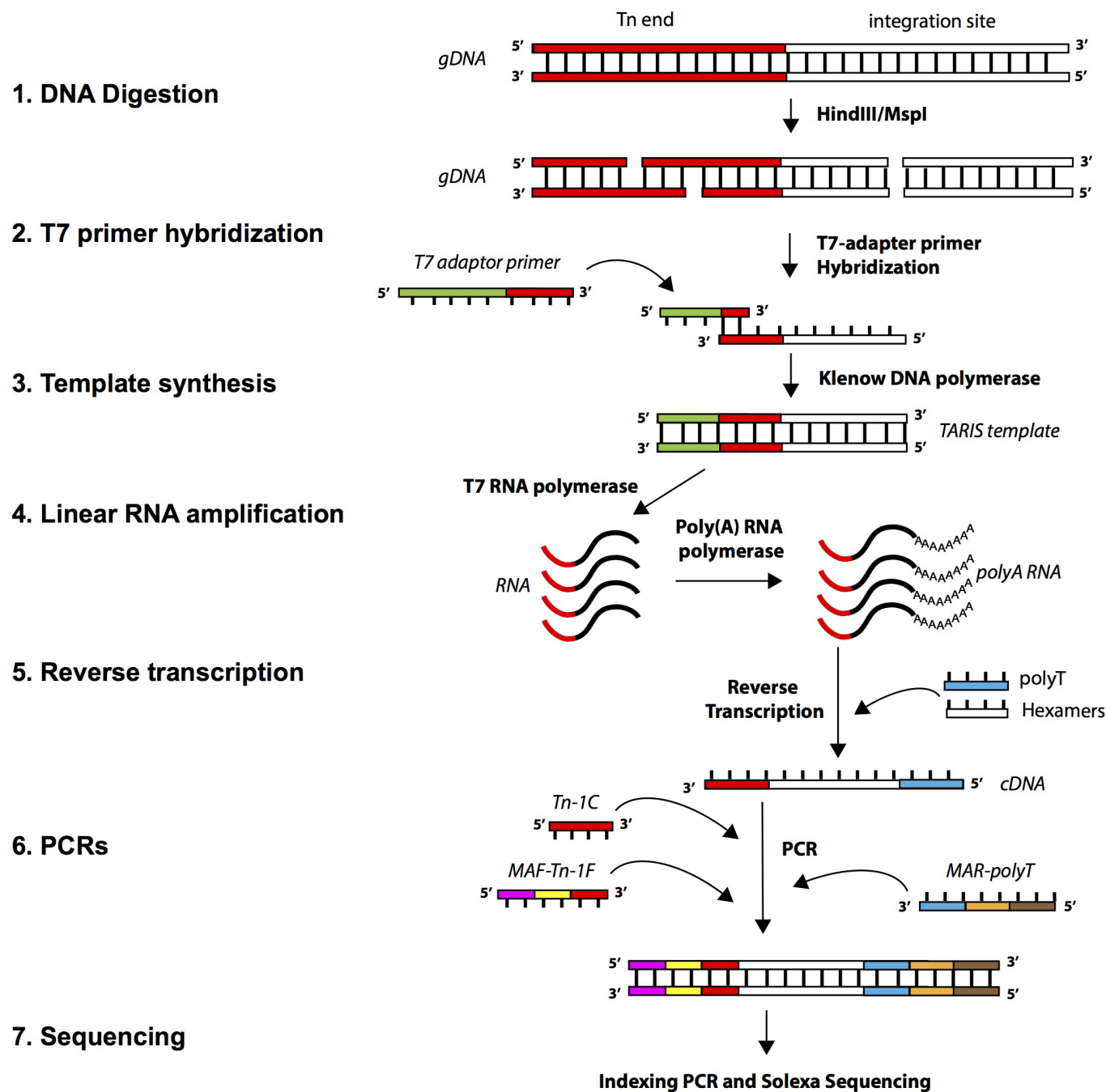
Data visualization using SPRING. We combined mRNA count matrices from five simultaneously processed and indexed libraries (LTHSC-2A, STHSC-2A, MPP4-2A, MPP3-2A, MPP2-2A). Cells with few mRNA counts ($< 1,000$ unique molecular identifiers) and stressed cells (mitochondrial gene-set Z-score > 1) were filtered out³⁵. The remaining high-quality cells (4,248) were total-counts normalized. We next filtered genes, keeping those that were well detected (mean expression > 0.05) and highly variable (CV > 2). Finally, we reduced dimensionality by Z-scoring each gene and applying principal component analysis (PCA), retaining the top 50 principal components. The cells were then visualized using SPRING, a graph-based single-cell viewing interface³⁶. Visual inspection of the SPRING plot revealed a strong cell-cycle signature defined by high expression of genes associated with the G2/M phase (*Ccnb1*, *Plk1*, *Cdc20*, *Aurka*, *Cenpf*, *Cenpa*, *Ccnb2*, *Birc5*, *Bub1*, *Bub1b*, *Ccnb2*, *Cks2*, *E2f5*, *Cdkn2d*). Hypothesizing that this

cell-cycle signature could affect high-dimensional distances between cells in a way that obscures their segregation by lineage-specific genes, we attempted to remove it³⁷. Specifically, we filtered from the analysis genes that were significantly correlated with the sum Z-score of G2/M genes ($P < 10^{-4}$, Bonferroni corrected; 401 genes total, resulting in 28,205 remaining genes). PCA and clustering analysis were repeated using the reduced gene list.

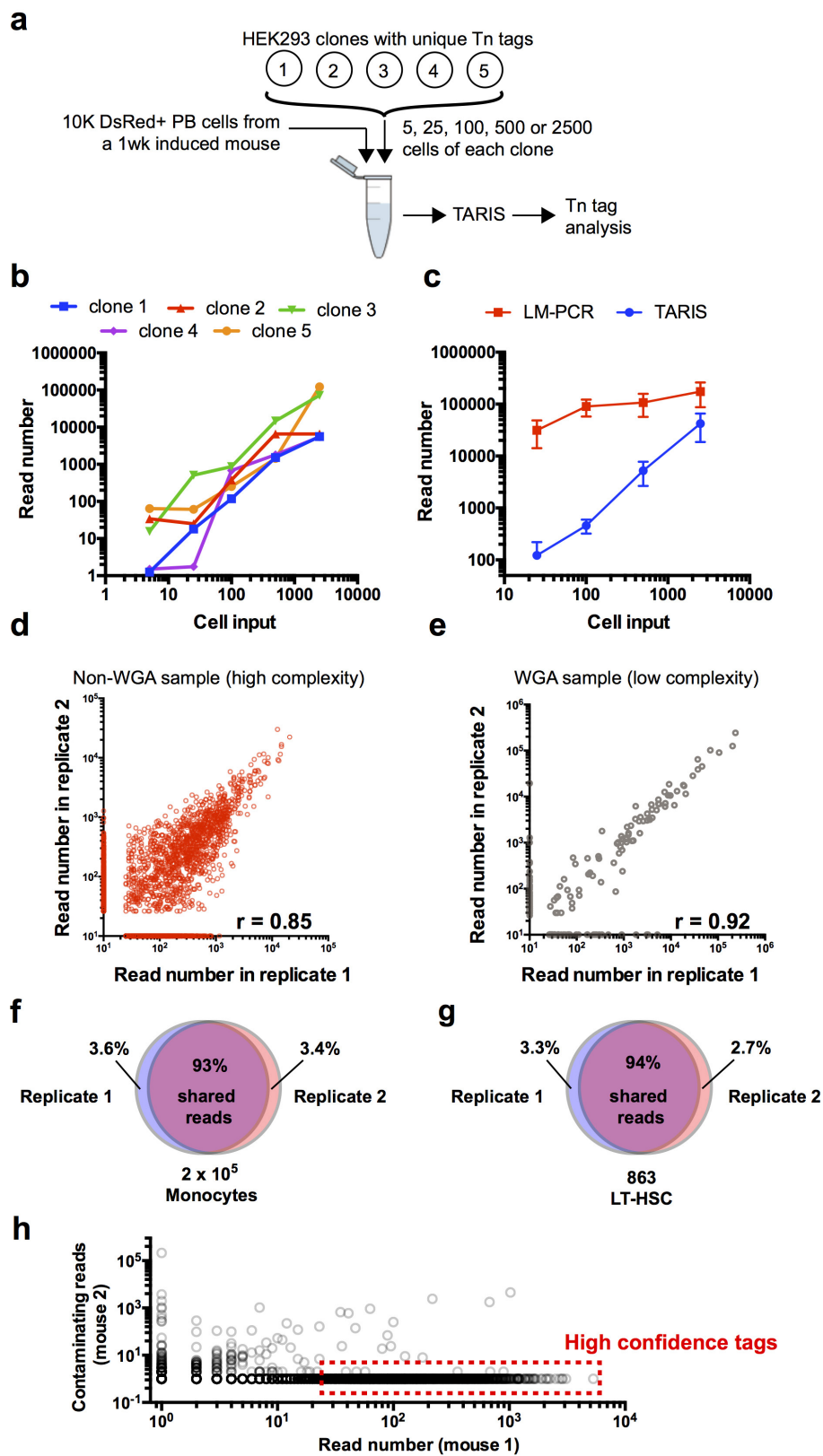
Clustering of single-cell profiles. We performed unsupervised clustering of the processed single-cell data with the Louvain–Jaccard method package from ref. 38. To assess cluster stability and choose the value of k , we downsampled 85% of cells and applied the Louvain–Jaccard method using 50 principal components. We tested k values from 10 to 30 and for each k we compared 100 times the randomly downsampled clustering using the Jaccard-index measurement in the R package fpc (Flexible Procedures for Clustering). We considered a Jaccard-index minimum of 0.75 as sufficiently robust and selected values of $k > 30$, which resulted in the identification of 11–12 clusters³⁹. Differential expression analysis was performed using the method package from ref. 38 (results are included in Supplementary Table 2).

Data availability. The Gene Expression Omnibus accession number is GSE90742. Additional data files will be made available upon reasonable request from the corresponding author. SPRING plots (with and without removal of the G2/M cell-cycle signature) are available for inspection at the following links: https://kleintools.hms.harvard.edu/tools/springViewer.html?cgi-bin/client_datasets/ARF2017_combined_nocycle and https://kleintools.hms.harvard.edu/tools/springViewer.html?cgi-bin/client_datasets/ARF2017_combined.

29. Foudi, A. *et al.* Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells. *Nat. Biotechnol.* **27**, 84–90 (2009).
30. Oguro, H., Ding, L. & Morrison, S. J. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* **13**, 102–116 (2013).
31. Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135**, 1118–1129 (2008).
32. Harkey, M. A. *et al.* Multiarm high-throughput integration site detection: limitations of LAM-PCR technology and optimization for clonal analysis. *Stem Cells Dev.* **16**, 381–392 (2007).
33. Wang, G. P. *et al.* DNA bar coding and pyrosequencing to analyze adverse events in therapeutic gene transfer. *Nucleic Acids Res.* **36**, e49 (2008).
34. Zilionis, R. *et al.* Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protocols* **12**, 44–73 (2017).
35. Ilicic, T. *et al.* Classification of low quality cells from single-cell RNA-seq data. *Genome Biol.* **17**, 29 (2016).
36. Weinreb, C., Wolock, S. & Klein, A. SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. Preprint at <https://doi.org/10.1101/090332> (2016).
37. Buettner, F. *et al.* Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol.* **33**, 155–160 (2015).
38. Shekhar, K. *et al.* Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell* **166**, 1308–1323 (2016).
39. Hennig, C. Cluster validation by measurement of clustering characteristics relevant to the user. Preprint at <http://arXiv:1703.09282> (2017).



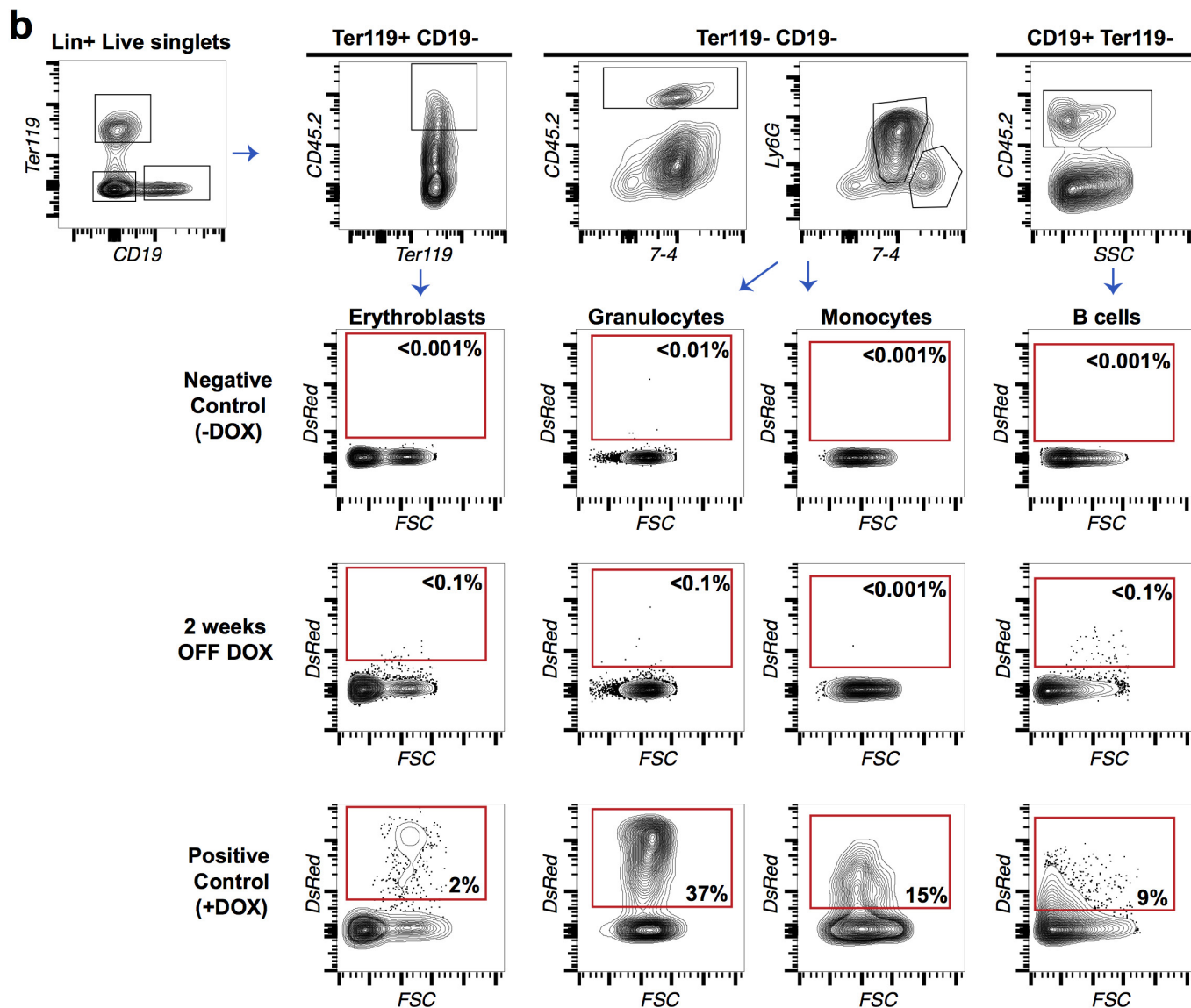
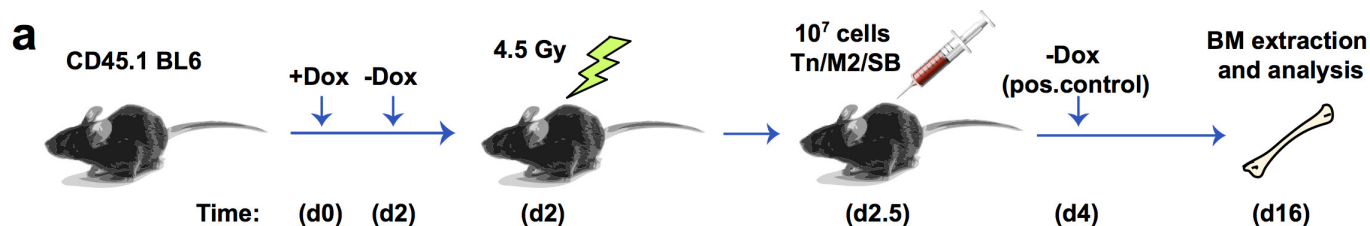
Extended Data Figure 1 | TARIS. Illustration of the TARIS procedure. The procedure is described in detail in the Methods.



Extended Data Figure 2 | See next page for caption.

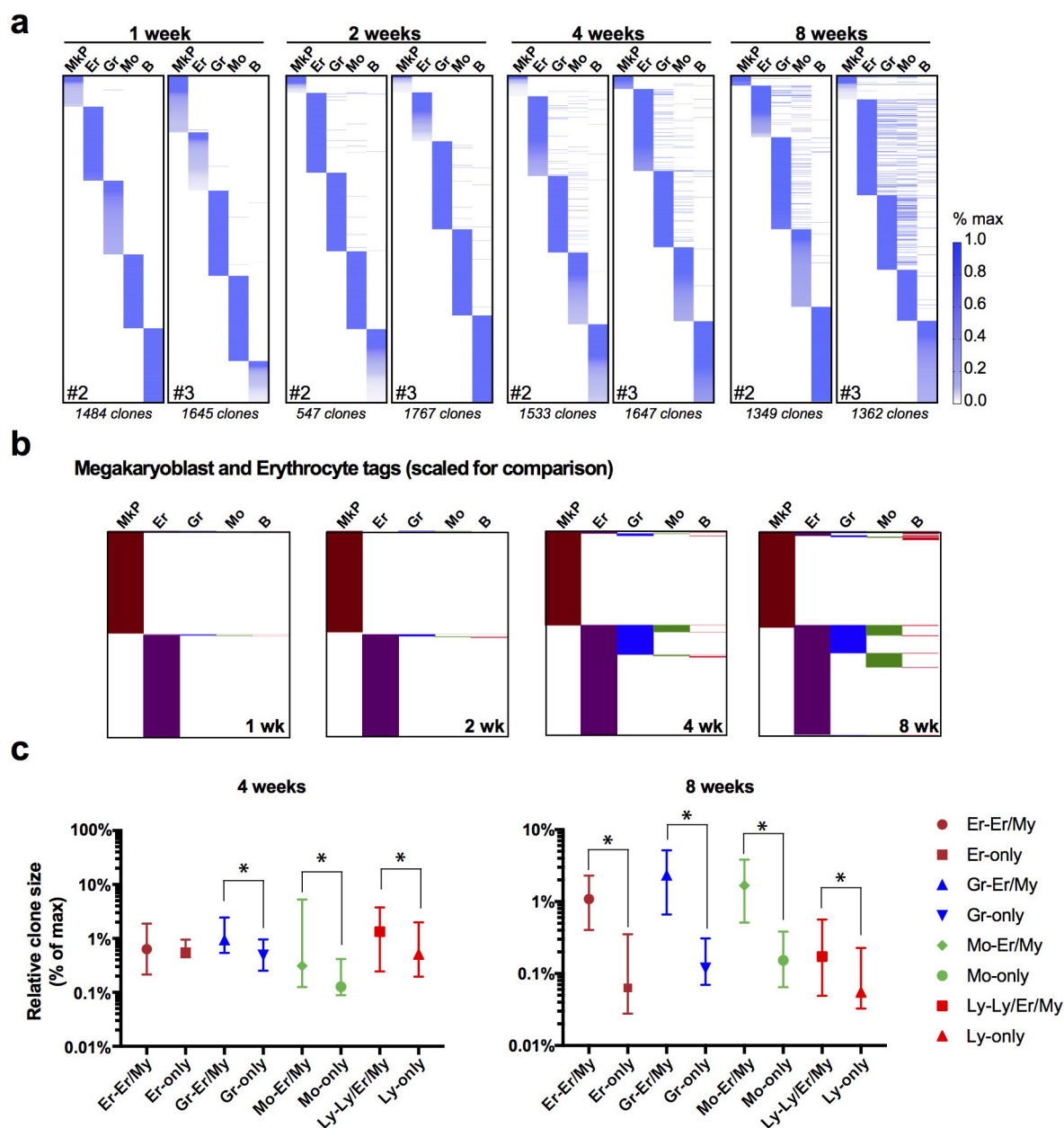
Extended Data Figure 2 | Evaluation of the TARIS method. **a**, Design for the detection limit experiment. Spike-ins of a known number of HEK293 cells carrying unique transposon integration tags were used in a mix of 10,000 DsRed⁺ peripheral blood cells from a freshly induced HSB mouse. **b**, Detection limit chart. Values represent the read number for each clone and for each number of input cells. Both axes are in log₁₀ scale. Values represent the sum of two independent experiments. **c**, Comparison of the average read number value between TARIS and the LM-PCR method. Values represent mean \pm s.d. of five different transposon tag clones. **d**, Reproducibility analysis in a non-whole-genome amplified sample with high complexity (2×10^5 bone marrow granulocytes 4 weeks after pulse). **e**, Reproducibility in a whole-genome amplified sample with low

complexity (863 LT-HSCs 4 weeks after pulse). **f**, Venn diagram showing overlapping transposon tag reads between two TARIS replicates from the same sample high-complexity sample (2×10^5 bone marrow monocytes at 4 weeks after induction). **g**, Venn diagram showing overlapping transposon tag reads between two TARIS replicates from the same low-complexity sample (863 LT-HSCs at 4 weeks after induction). **h**, Contamination analysis between samples from two different mice. The plot represents the read numbers of tags from Lin⁺ populations from mouse 1, and their read number values in Lin⁺ populations in mouse 2. High-confidence tags are selected as those tags with more than 25 reads, and at least 10 times higher read count compared with any of the samples from a separate mouse.



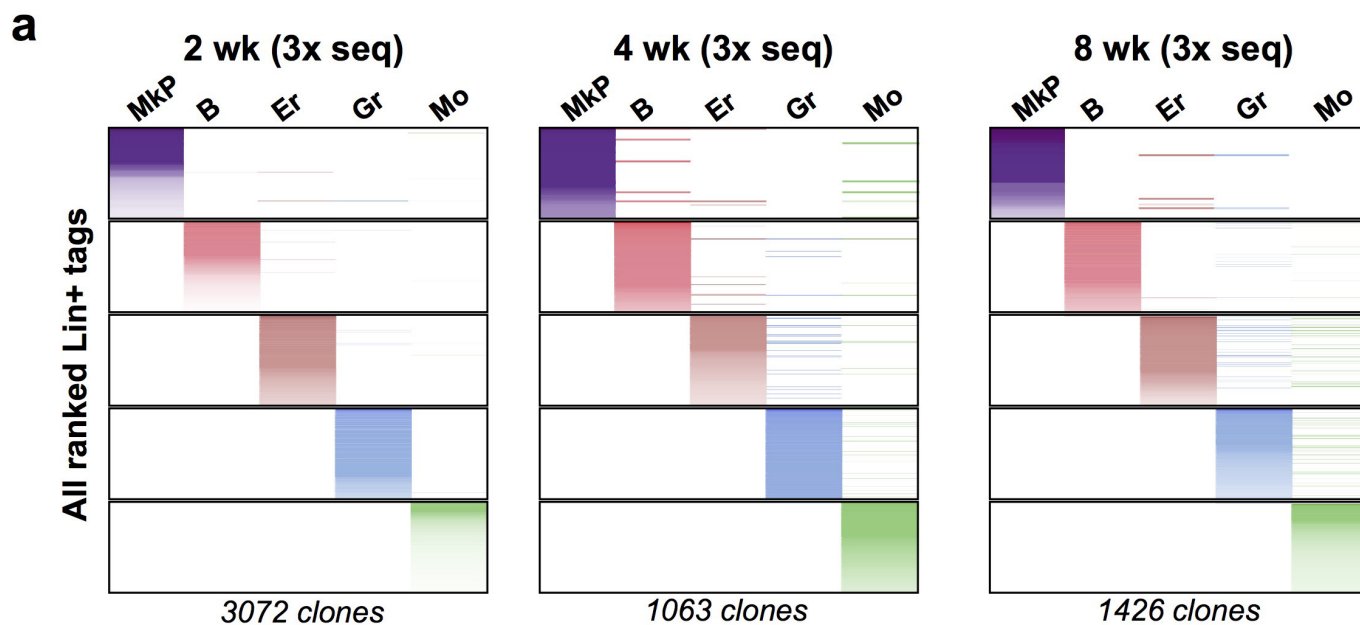
Extended Data Figure 3 | Analysis of residual HSB activity after Dox withdrawal. **a**, Experimental design. Residual HSB activity after Dox removal was assayed by transplantation into CD45.1 mice. Sub-lethally irradiated recipients were treated with Dox for 48 h. Dox was removed 12 h before transplantation. Ten million whole bone marrow cells were transplanted and mice were allowed to recover for 2 weeks. As a positive

control, mice were continuously treated with Dox until 48 h after transplant. As a negative control, cells were transplanted into non-Dox treated mice. DsRed labelling was analysed as a proxy for HSB activity in granulocytes, erythroblasts, monocytes, and B cells. **b**, FACS plots showing the negligible labelling of CD45.2 M2/HSB/Tn cells in transplanted recipients 24 h after Dox removal.



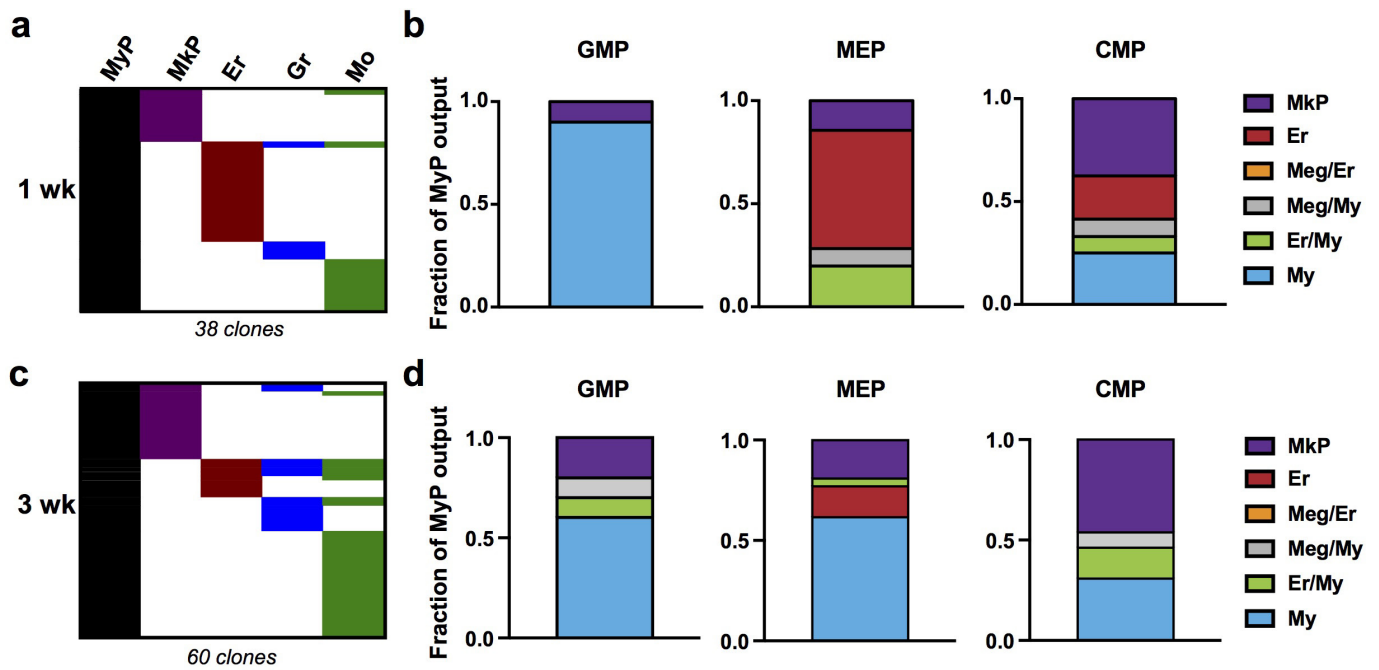
Extended Data Figure 4 | Additional representations and analyses of Lin⁺ tags. **a**, Lin⁺ clones of the second and third mice used for quantifications in Fig. 1d–g. **b**, Scale-adjusted binary (presence/absence) representation of all detected MkP and erythroblast tags in the

experiments from Fig. 1d–g. **c**, Relative quantification of scale-normalized clone sizes for each lineage, comparing unilineage versus oligo/multilineage clones. Values are interquartile range and median from three independent mice at four weeks and eight weeks after induction.



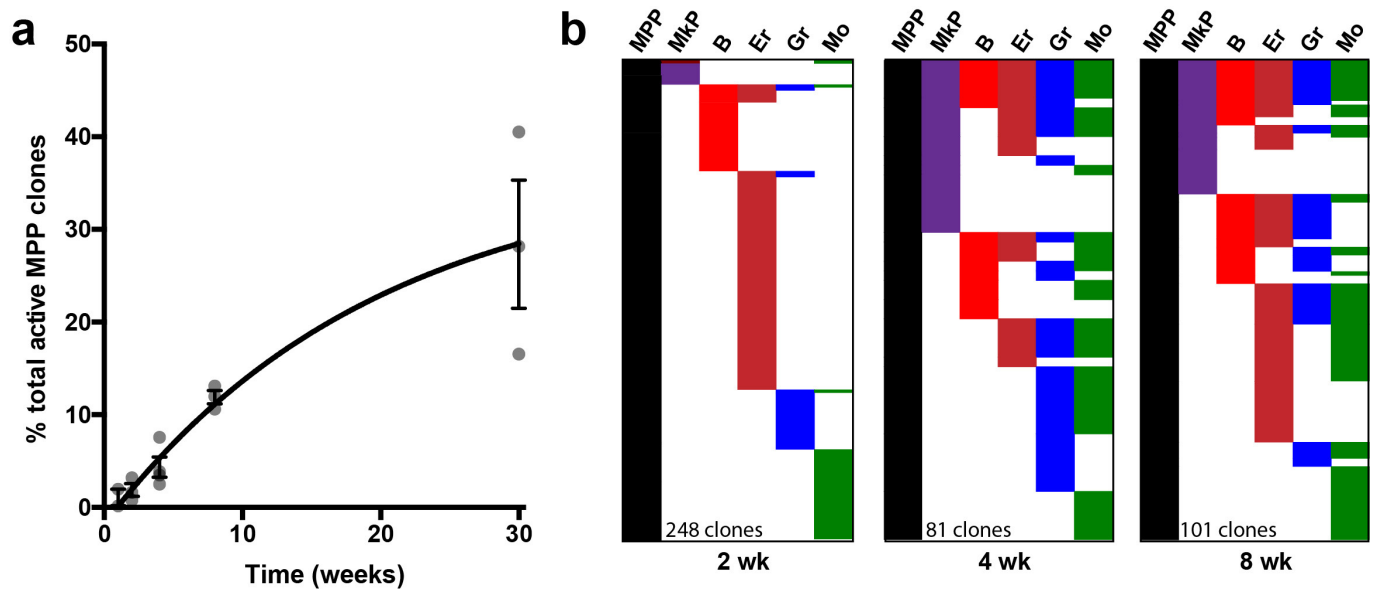
Extended Data Figure 5 | Validation of granulocyte/monocyte/erythroblast and megakaryocyte-restricted tags. **a**, Three independent transposon tag libraries were prepared and sequenced from 2-week-,

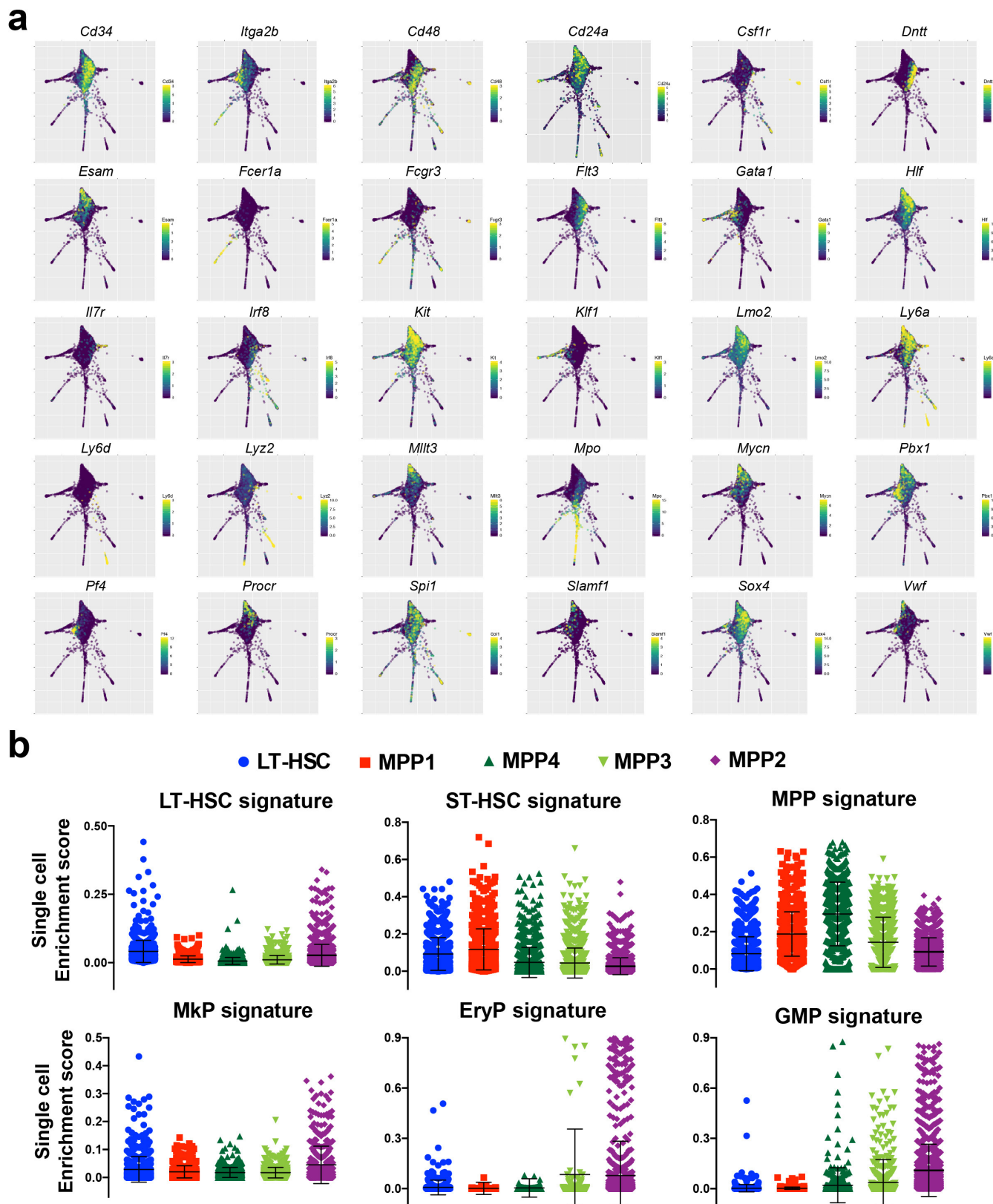
4-week-, and 8-week-chased mice. Reads from the three libraries were then pooled together for each lineage.



Extended Data Figure 6 | Lineage fate of myeloid progenitors. **a**, Two M2/HSB/Tn mice were induced and chased for one week, and then myeloid progenitors (GMP, MEP, and CMP) and Lin⁺ cells were isolated from bone marrow and their transposon tag content was analysed. Chart is a binary representation of all Lin⁺ tags overlapping with any myeloid progenitor tag ranked by lineage. **b**, Quantification of relative lineage

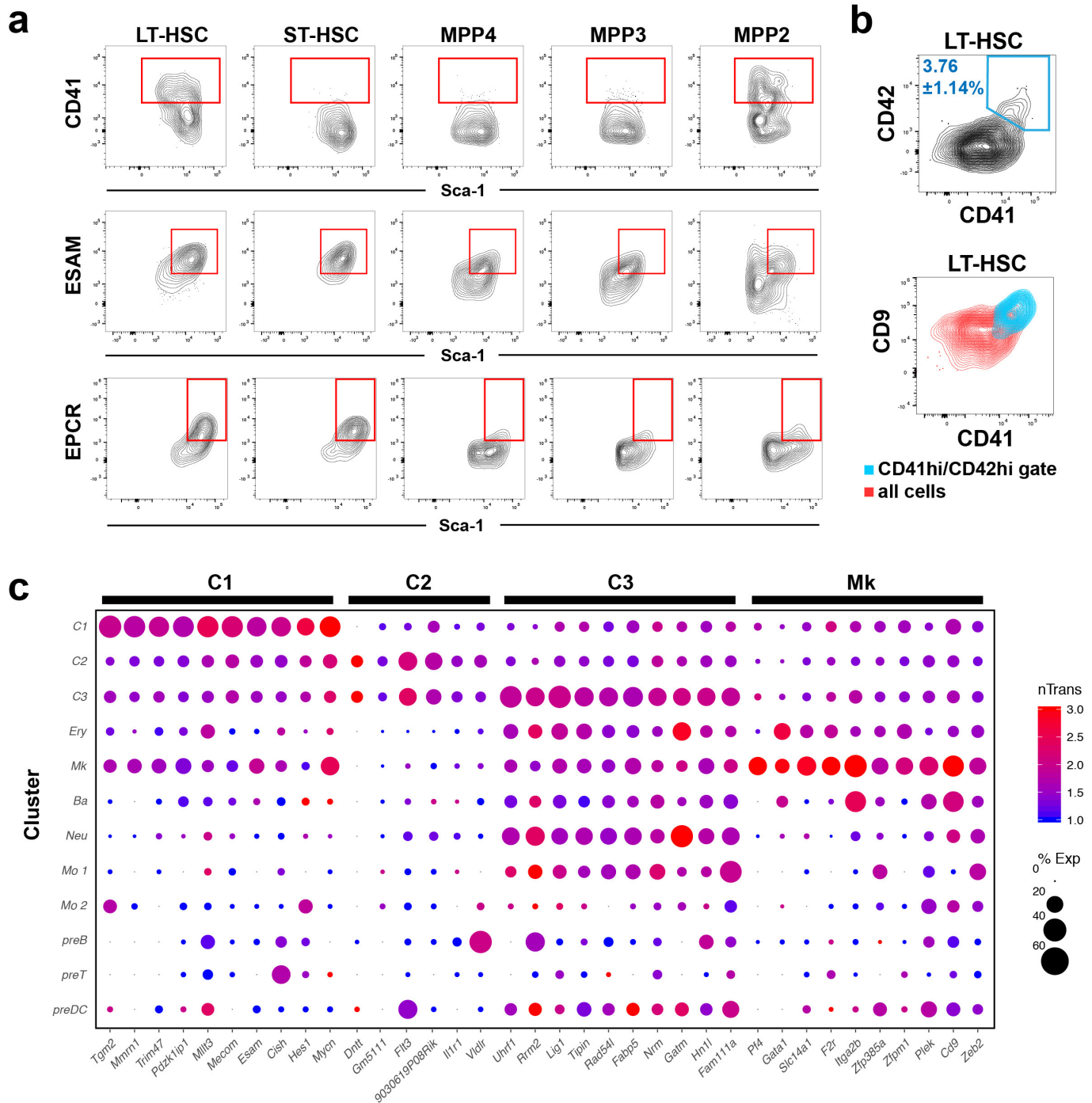
contribution of GMPs, MEPs, and CMPs as a fraction of lineage-specific/total lineage-overlapping clones for each MyP subset. Values are mean of the two analysed mice. **c**, An additional M2/HSB/Tn mouse was induced and chased for three weeks, and then processed as in **a**. **d**, Quantification of relative lineage contribution of GMPs, MEPs, and CMPs at three weeks after labelling.





Extended Data Figure 8 | Single-cell heterogeneity of HSC/MPPs.
a, SPRING plots showing selected differentially expressed markers. Scale represents amount of detected mRNA copies (normalized) of each marker gene. **b**, Enrichment score analysis for single cells in each FACS-sorted

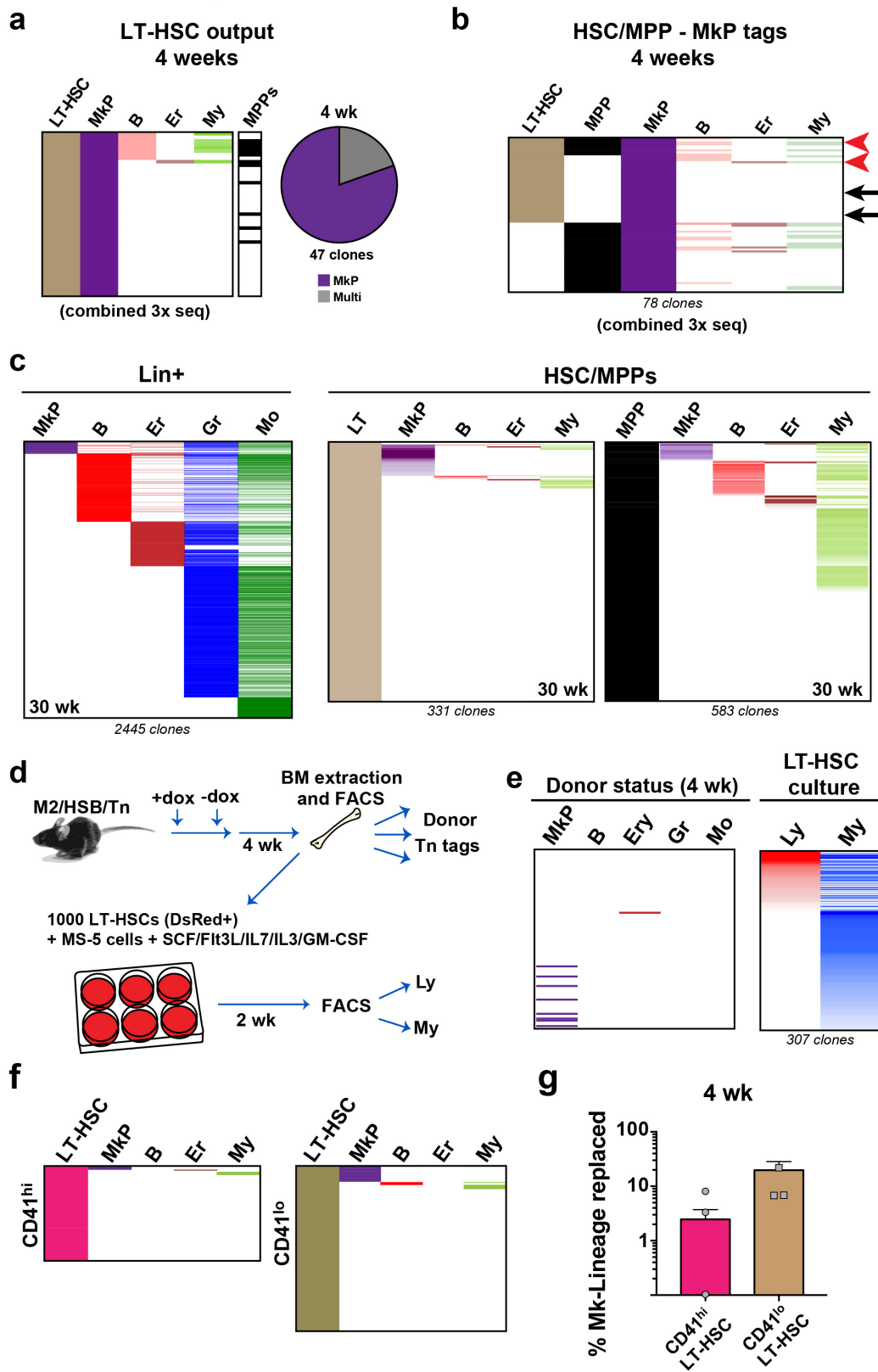
population compared with previously obtained bulk transcriptional signatures of bone marrow populations sorted using traditional markers (from the Immgen database).



Differentially expressed genes

Extended Data Figure 9 | Differentially expressed markers for clusters C1, C2, C3, and megakaryocyte. **a**, FACS plots showing heterogeneity in expression of cluster markers within the analysed HSC/MPP subsets. **b**, FACS plots showing expression of different megakaryocyte-primed

cluster markers (CD41, CD42, and CD9) within the LT-HSC gate. **c**, The expression value (nTrans) and percentage of expressing cells of each cluster (% Exp). The top ten highest expressed genes that distinguish each cluster are shown.



Extended Data Figure 10 | See next page for caption.

Extended Data Figure 10 | Additional data on clonal origin of MkP.

a, Three independent transposon tag libraries were prepared and sequenced for LT-HSC, MPP, and the five Lin⁺ populations, from one mouse at four weeks. Each column represents the combined tags detected from three amplicon libraries prepared for each population, to facilitate visualization of the smallest clones. Tags are coloured by frequency in each lineage, and organized by rank. **b**, Origin of megakaryocytes. Alignment of all MkP clones that had detectable tags in HSC/MPPs from a mixed library combining three independent sequencing reactions. Tags are coloured by frequency in each lineage (except for MkP), and organized by rank. Arrows indicate tags verified by clone-specific PCR. **c**, Alignment of

transposon tags from all Lin⁺ populations, LT-HSCs, and MPPs collected from 30-week-chased mice. Tags are coloured by frequency in each lineage, and organized by rank. **d**, Experimental design for testing *in vitro* myeloid and lymphoid potential from sorted LT-HSCs. **e**, *In vitro* myeloid potential of LT-HSCs. Alignment of donor Lin⁺ tags with transposon tags obtained from myeloid and lymphoid cells derived from donor LT-HSCs after two weeks in culture. **f**, Clonal output of CD41^{hi} and CD41^{lo} LT-HSCs at four weeks after labelling. **g**, Quantification of megakaryocyte lineage replacement by CD41^{hi} versus CD41^{lo} LT-HSCs (measured as the percentage of overlapping/total MkP reads) at four weeks after labelling. Values are mean \pm s.e.m. of three independent mice.

Treatment of autosomal dominant hearing loss by *in vivo* delivery of genome editing agents

Xue Gao^{1,2,3,†*}, Yong Tao^{4,5,†*}, Veronica Lamas⁴, Mingqian Huang⁴, Wei-Hsi Yeh^{1,2,3,6}, Bifeng Pan⁷, Yu-Juan Hu^{4,5}, Johnny H. Hu^{1,2,3}, David B. Thompson^{1,2}, Yilai Shu^{4,8}, Yamin Li⁹, Hongyang Wang^{4,10}, Shiming Yang¹⁰, Qiaobing Xu⁹, Daniel B. Polley⁴, M. Charles Liberman⁴, Wei-Jia Kong⁵, Jeffrey R. Holt⁷, Zheng-Yi Chen^{4§} & David R. Liu^{1,2,3§}

Although genetic factors contribute to almost half of all cases of deafness, treatment options for genetic deafness are limited^{1–5}. We developed a genome-editing approach to target a dominantly inherited form of genetic deafness. Here we show that cationic lipid-mediated *in vivo* delivery of Cas9–guide RNA complexes can ameliorate hearing loss in a mouse model of human genetic deafness. We designed and validated, both *in vitro* and in primary fibroblasts, genome editing agents that preferentially disrupt the dominant deafness-associated allele in the *Tmc1* (transmembrane channel-like gene family 1) Beethoven (*Bth*) mouse model, even though the mutant *Tmc1*^{*Bth*} allele differs from the wild-type allele at only a single base pair. Injection of Cas9–guide RNA–lipid complexes targeting the *Tmc1*^{*Bth*} allele into the cochlea of neonatal *Tmc1*^{*Bth/+*} mice substantially reduced progressive hearing loss. We observed higher hair cell survival rates and lower auditory brainstem response thresholds in injected ears than in uninjected ears or ears injected with control complexes that targeted an unrelated gene. Enhanced acoustic startle responses were observed among injected compared to uninjected *Tmc1*^{*Bth/+*} mice. These findings suggest that protein–RNA complex delivery of target gene-disrupting agents *in vivo* is a potential strategy for the treatment of some types of autosomal-dominant hearing loss.

Although about 100 deafness-associated alleles have been identified, few treatments are available to slow or reverse genetic deafness^{4,5}. Complementation of wild-type alleles, or silencing of dominant-negative mutant alleles, have shown promising results in animal models^{6,7}. Nonetheless, current approaches face potential challenges including immunogenicity, oncogenicity, and limitations of viral vectors^{8,9}.

Cas9-based genome editing agents can mediate targeted gene disruption or repair^{10–13}. For applications that seek a one-time, permanent modification of genomic DNA, the delivery of non-replicable, transient Cas9–single guide RNA (sgRNA) ribonucleotide protein (RNP) complexes *in vivo* offers improved DNA specificity and potentially greater safety and applicability^{14,15}, compared with methods that introduce DNA expressing these agents. Approximately 20% of alleles associated with genetic deafness are dominantly inherited¹. As Cas9–sgRNA complexes can efficiently disrupt genes through end-joining processes, we sought to design Cas9–sgRNA complexes that selectively disrupt dominant alleles associated with hearing loss.

Many genes linked to genetic hearing loss affect the function of sensory hair cells, which transduce acoustic vibrations into electrical

nerve signals. TMC1 is an essential component of mechanotransduction channels in mammalian hair cells¹⁶. Mutations in *TMC1* have been linked to recessive and dominant genetic deafness in humans¹⁷. A dominant-negative missense mutation in *TMC1* (p.M418K, c.T1253A) causes reduced single-channel current levels and calcium permeability¹⁶ in hair cells, and progressive post-lingual sensorineural hearing loss in humans^{18–20}. The *Tmc1*^{*Bth/+*} mouse model carries the orthologous missense mutation (p.M412K, c.T1235A) in the mouse *Tmc1* gene and exhibits progressive elevation of the auditory response threshold and progressive hair cell loss beginning at one month of age²¹. As the orthologous mutations in human and mouse both cause progressive, profound hearing loss, the *Tmc1*^{*Bth/+*} mouse is a promising model for the development of treatment strategies²¹.

We began by developing a genome editing strategy that preferentially disrupts the mouse mutant *Tmc1*^{*Bth*} allele. To distinguish the mutant and wild-type alleles, we identified sgRNAs that target *Tmc1* at sites that include the T1235A mutation and a nearby NGG protospacer-adjacent motif (PAM) sequence required by *Streptococcus pyogenes* Cas9. We identified three candidate sgRNAs (*Tmc1*-mut1, *Tmc1*-mut2 and *Tmc1*-mut3) that place the *Bth* mutation at position 11, 12, and 15, respectively, of the spacer, counting the PAM as positions 21–23 (Fig. 1a). Mismatches between the sgRNA and genomic DNA that are close to the PAM are poorly tolerated by Cas9¹⁰, increasing the likelihood that the *Bth* mutant allele will be selectively edited. A fourth sgRNA, *Tmc1*-mut4, is a truncated version of *Tmc1*-mut3 designed to increase genome editing DNA specificity²². We evaluated the ability of these four sgRNAs in complex with Cas9 to cleave either the wild-type *Tmc1* or the *Tmc1*^{*Bth*} allele *in vitro*. All sgRNAs tested comparably or preferentially cleaved the *Tmc1*^{*Bth*} allele, with *Tmc1*-mut3 exhibiting the greatest selectivity (Extended Data Fig. 1a, b).

We performed lipid-mediated delivery of Cas9–sgRNA RNP complexes into cultured primary fibroblasts derived from wild-type or homozygous *Tmc1*^{*Bth/Bth*} mice to evaluate the allele specificity of genomic DNA modification in mouse cells. We delivered Cas9 complexed with each of the four sgRNAs using Lipofectamine 2000 into both wild-type and *Tmc1*^{*Bth/Bth*} mutant fibroblasts. RNP delivery into these primary fibroblasts was twofold to fourfold less efficient than with HEK293T cells (Extended Data Fig. 1c). The highest rate of targeted insertions and deletions (indels) in mutant *Tmc1*^{*Bth/Bth*} fibroblasts (10%) was observed with Cas9–*Tmc1*-mut3 RNPs, while lower indel frequencies (0.74–4.1%) were observed using the other sgRNAs (Fig. 1b).

¹Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA. ²Howard Hughes Medical Institute, Harvard University, Cambridge, Massachusetts 02138, USA. ³Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02141, USA. ⁴Department of Otolaryngology and Program in Neuroscience, Harvard Medical School and Eaton Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts 02114, USA. ⁵Department of Otorhinolaryngology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430022, China. ⁶Program in Speech and Hearing Bioscience and Technology, Harvard University, Cambridge, Massachusetts 02138, USA. ⁷Departments of Otolaryngology and Neurology, F.M. Kirby Neurobiology Center Boston Children's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. ⁸Department of Otolaryngology–Head and Neck Surgery, Eye and ENT Hospital, Shanghai Medical College, Fudan University, Shanghai, China. ⁹Department of Biomedical Engineering, Tufts University, Medford, Massachusetts 02155, USA. ¹⁰Department of Otolaryngology & Head Neck Surgery, Key Lab of Hearing Impairment Science of Ministry of Education, Key Lab of Hearing Impairment Prevention and Treatment of Beijing City, Chinese PLA Medical School, Beijing, China. [†]Present addresses: Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas 77005, USA (X.G.); Department of Otolaryngology–Head and Neck Surgery, Shanghai Ninth People's Hospital and Ear Institute, Shanghai Jiaotong University School of Medicine, Shanghai, 200011, China (Y.T.).

*These authors contributed equally to this work.

§These authors jointly supervised this work.

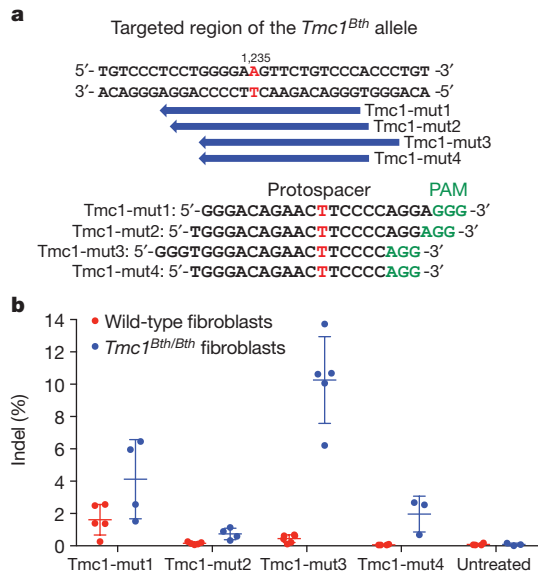


Figure 1 | Design of a genome-editing strategy to disrupt the *Tmc1^{Bth}* mutant allele. **a, SpCas9 sgRNAs were designed to target the mutant *Tmc1^{Bth}* allele, in which T1235 is changed to A (red). The protospacer (blue arrows) of each *Tmc1^{Bth}*-targeting sgRNA contains a complementary T (red) that pairs with the T1235A mutation in the *Tmc1^{Bth}* allele, but that forms a mismatch with the wild-type *Tmc1* allele. **b**, Lipid-mediated delivery of Cas9–sgRNA complexes into primary fibroblasts derived from wild-type or homozygous *Tmc1^{Bth/Bth}* mice. Purified Cas9 protein (100 nM) and 100 nM of each sgRNA shown were delivered using Lipofectamine 2000. Indels were quantified by HTS. Individual values ($n = 3–6$) are shown; horizontal lines and error bars represent the mean values \pm s.d. of three or more independent biological replicates.**

By contrast, all tested sgRNAs edited the wild-type *Tmc1* locus much less efficiently in wild-type fibroblasts (0.066–1.6% indels) (Fig. 1b). Notably, Cas9–Tmc1-mut3 modified the mutant *Tmc1^{Bth}* allele 23-fold more efficiently than the wild-type allele (Fig. 1b and Extended Data Fig. 1d). We also prepared three corresponding wild-type *Tmc1*-targeting sgRNAs (Tmc1-wt1, Tmc1-wt2 and Tmc1-wt3) that lack the T1235A mutation. These sgRNAs edited wild-type fibroblasts on average tenfold more efficiently than *Tmc1^{Bth/Bth}* fibroblasts (Extended Data Fig. 1e), confirming that the observed allele selectivities did not arise from the inability of the wild-type *Tmc1* allele to be edited.

We tested 17 cationic lipids for their ability to deliver the Cas9–Tmc1-mut3 RNP into *Tmc1^{Bth/Bth}* fibroblasts. Several lipids supported substantial modification of the target locus, including RNAiMAX (7.7%), CRISPRMAX (8.9%), and Lipofectamine 2000 (12%) (Extended Data Fig. 2). By contrast, treatment of wild-type fibroblasts with Cas9–Tmc1-mut3 and the same set of 17 lipids resulted in low ($\leq 0.5\%$) indel rates (Extended Data Fig. 2a, b). These results suggest that the target mutant *Tmc1* locus can be preferentially disrupted by Cas9–guide RNA complexes.

Exposure of cells to Cas9–sgRNA agents typically results in the modification of both on-target and off-target loci^{10,23}. We used both the GUIDE-seq method²³ and computational prediction²⁴ to identify potential off-target loci that could be modified by exposure to Cas9–Tmc1-mut3. Ten off-target sites containing up to six mismatches in the protospacer region of Tmc1-mut3 sgRNA were identified by GUIDE-seq (Extended Data Fig. 3a). None of these off-target loci are known to be associated with hearing function (Extended Data Table 1a). We measured the indel frequency at each off-target site by high-throughput sequencing (HTS) in *Tmc1^{Bth/Bth}* primary mouse fibroblasts treated with Cas9–Tmc1-mut3 following plasmid DNA nucleofection or RNP delivery. Plasmid nucleofection resulted in 0.68–8.1% indels at nine of the ten GUIDE-seq-identified off-target sites (Extended Data Fig. 3b and Extended Data Table 1a). By contrast, after RNP delivery,

modification of only one off-target site (*off-T1*, 1.2% indels) was detected (Extended Data Fig. 3b), consistent with our earlier findings that RNP delivery greatly reduces off-target editing compared with DNA delivery¹⁵. Among the computationally predicted off-target sites²⁴, only the two (*off-T1'* and *off-T2'*) that were also identified as off-targets by GUIDE-seq were observed to undergo modification (Extended Data Table 1b). Together, these results suggest that delivery of Cas9–Tmc1-mut3 RNP complexes into *Tmc1^{Bth/Bth}* cells leads to minimal off-target modification, and that phenotypes affecting hearing are unlikely to arise from off-target modification.

To evaluate the ability of the Cas9–Tmc1-mut3 sgRNA complex to target the *Tmc1^{Bth}* allele in hair cells *in vivo*, we complexed Cas9–Tmc1-mut3 sgRNA with Lipofectamine 2000 and injected the resulting mixture into the scala media of neonatal mice by cochleostomy. Neonatal cochlear hair cells produce TMC1 and TMC2, both of which can enable sensory transduction. To isolate the effect of editing the *Tmc1^{Bth}* allele, we injected the Cas9–Tmc1-mut3 sgRNA–lipid complex into *Tmc1^{Bth/Δ}Tmc2^{Δ/Δ}* mice¹⁶, to avoid transduction current contributions from TMC2 or wild-type TMC1. We recorded sensory transduction currents from inner hair cells (IHCs) after injection with the Cas9–Tmc1-mut3–lipid complex, or with a control targeting an unrelated *Gfp* gene. We observed a significant decline in transduction current amplitudes in *Tmc1^{Bth/Δ}Tmc2^{Δ/Δ}* mice following injection with Cas9–Tmc1-mut3–lipid complexes, consistent with disruption of the *Tmc1^{Bth}* allele in sensory hair cells *in vivo*, but not with Cas9–GFP sgRNA–lipid complexes (Fig. 2a, b).

In *Tmc1^{Bth/+}Tmc2^{+/+}* mice (referred to hereafter as *Tmc1^{Bth/+}* mice), IHCs undergo progressive death, followed by the outer hair cells (OHCs)²¹. To determine the effect of Cas9–Tmc1-mut3 sgRNA on *Tmc1^{Bth/+}* hair cell survival, we injected Cas9–Tmc1-mut3–lipid complex into the scala media of mice on postnatal day 1 (P1) and removed the injected and uninjected cochleae after eight weeks. Uninjected ears exhibited substantial loss of IHCs and partial degeneration of OHCs (Fig. 2c, f, g) compared with wild-type ears (C3HeB/FeJ (C3H) mice, which are the genetic background of the *Tmc1^{Bth/+}* mice) (Fig. 2e). In injected ears, survival of IHCs and OHCs was significantly enhanced (Fig. 2d, f, g). Stereocilia bundles were observed on surviving IHCs in injected ears, but were absent in uninjected ears in the basal and middle turns (Extended Data Fig. 4a). These results suggest that Cas9–Tmc1-mut3–lipid injection *in vivo* promotes hair cell survival in *Tmc1^{Bth/+}* mice. The strong differences between treated and untreated ears suggests that sporadic disruption of *Tmc1^{Bth}* may benefit not only edited hair cells, but also surrounding hair cells, consistent with previous findings²⁵.

To study the effect of Cas9–Tmc1-mut3–lipid injection on cochlear function in *Tmc1^{Bth/+}* mice, we measured auditory brainstem responses (ABRs), which represent the sound-evoked neural output of the cochlea, as well as distortion product otoacoustic emissions (DPOAEs), which measure the amplification provided by OHCs²¹. In uninjected ears of *Tmc1^{Bth/+}* mice, we observed profound attenuation of cochlear neural responses, with ABR thresholds ranging from 70–90 dB at four weeks of age, compared with 30–50 dB for wild-type C3H mice (Fig. 3a and Extended Data Fig. 4b). The elevations in DPOAE thresholds at this time in *Tmc1^{Bth/+}* mice were smaller than the elevations in ABR threshold (Extended Data Fig. 5a), consistent with reports that IHCs are more severely affected than OHCs in *Tmc1^{Bth/+}* mice²¹. Four weeks after Cas9–Tmc1-mut3–lipid injection, treated *Tmc1^{Bth/+}* ears showed substantially enhanced cochlear function, with lower ABR thresholds relative to uninjected ears at all frequencies below 45 kHz (Fig. 3a). Significant ($P \leq 0.001$) hearing preservation was detected from 8 to 23 kHz, with average ABR thresholds 15 dB lower for treated ears than untreated contralateral ears (Fig. 3a; Supplementary Table 1). DPOAE thresholds were slightly elevated in the injected ears, consistent with OHC damage, perhaps from the injection procedure (Extended Data Fig. 5). We also observed greater ABR wave 1 amplitudes, and a more normal

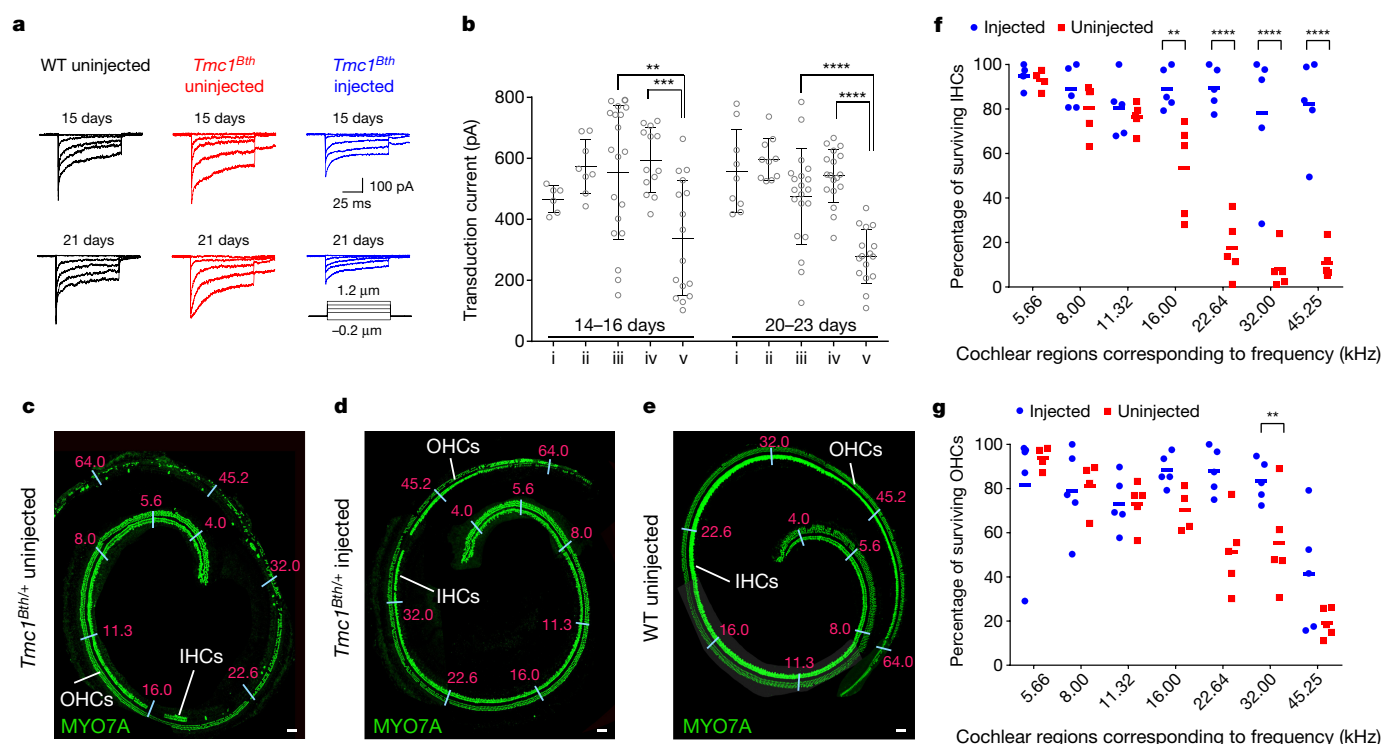


Figure 2 | Effects of Cas9-Tmc1-mut3 sgRNA-lipid injection on hair-cell function and hair-cell survival in mice. **a**, Representative transduction currents from IHCs of P0–P1 wild-type (WT) or *Tmc1*^{Bth/Δ}*Tmc2*^{Δ/Δ} mice that were uninjected, or injected with the Cas9-Tmc1-mut3-lipid complex, 15 or 21 days after injection. **b**, Maximal transduction current amplitudes for 135 IHCs from wild-type C57B/L6 and *Tmc1*^{Bth/Δ}*Tmc2*^{Δ/Δ} mice. i, uninjected wild-type C57B/L6 mice; ii, wild-type C57B/L6 mice injected at P1 with Cas9-Tmc1-wt3-lipid; iii, uninjected *Tmc1*^{Bth/Δ}*Tmc2*^{Δ/Δ} mice; iv, *Tmc1*^{Bth/Δ}*Tmc2*^{Δ/Δ} mice injected at P1 with Cas9-GFP sgRNA-lipid; v, *Tmc1*^{Bth/Δ}*Tmc2*^{Δ/Δ} mice injected at P1 with Cas9-Tmc1-mut3-lipid. Data were recorded after 14–23 days. Individual values ($n = 6–20$) are shown; horizontal lines and error bars reflect mean \pm s.d. **c–e**, Representative

confocal microscopy images around the age of eight weeks from an uninjected *Tmc1*^{Bth/+} cochlea (**c**); the contralateral cochlea of the mouse in **c** injected with Cas9-Tmc1-mut3-lipid complex at P1 (**d**); and an untreated wild-type C3H cochlea (**e**). Numbers in pink indicate approximate frequencies (in kHz) sensed by each region. Scale bars, 50 μ m. **f**, **g**, Quantification of IHC (**f**) and OHC (**g**) survival percentages in *Tmc1*^{Bth/+} mice relative to wild-type C3H mice (100%) eight weeks after Cas9-Tmc1-mut3-lipid injection (blue) compared to uninjected (red) contralateral ears. Individual values are shown; horizontal lines represent mean values of five biological replicates. Statistical tests in **b** are two-population *t*-tests, and in **f**, **g** are two-way ANOVAs with Bonferroni correction: ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

ABR waveform pattern, in injected ears than in uninjected controls (Fig. 3b, c). Together, these results show that injection of neonatal *Tmc1*^{Bth/+} mice with Cas9-Tmc1-mut3-lipid complexes reduces progressive hearing loss.

To test whether amelioration of hearing loss requires the mutant *Tmc1*^{Bth} allele-specific sgRNA, we injected Cas9-Tmc1-wt3-lipid complexes targeting the wild-type *Tmc1* allele rather than the *Tmc1*^{Bth} mutant allele into P1–2 *Tmc1*^{Bth/+} mice. After four weeks, ABR thresholds in the injected ears were similar to, or worse than, those in the contralateral uninjected ears (Extended Data Fig. 6a; Supplementary Table 1), consistent with the inability of Cas9-Tmc1-wt3 to efficiently disrupt the *Tmc1*^{Bth} allele (Extended Data Fig. 1e), and possible disruption of wild-type *Tmc1*. Injection of Cas9-sgRNA-lipid complexes targeting an unrelated gene (*Gfp*) did not significantly affect ABR thresholds at most tested frequencies in *Tmc1*^{Bth/+} mice (Extended Data Fig. 6b). To test whether preservation of cochlear function requires Cas9 nuclease activity, rather than transcriptional interference from Cas9 binding to *Tmc1*, we treated *Tmc1*^{Bth/+} mice with catalytically inactive dCas9¹⁰ complexed with Tmc1-mut1 sgRNA and observed no evidence of hearing preservation (Extended Data Figs 5d, 6c; Supplementary Table 1). To evaluate the effects of the treatment on normal mice, we injected Cas9-Tmc1-mut3-lipid into wild-type C3H mice. We observed similar or slightly elevated ABR thresholds in injected ears relative to uninjected ears four weeks after treatment (Extended Data Fig. 6d, e), suggesting that Cas9-Tmc1-mut3 does not modify wild-type *Tmc1* efficiently enough to substantially affect hearing. Finally, injection of Cas9 and lipid without sgRNA did not improve ABR or DPOAE thresholds

(Extended Data Fig. 6f, g). Collectively, these results establish that hearing preservation depends on sgRNA allele specificity, Cas9 DNA cleavage activity, and the presence of the *Tmc1*^{Bth} allele. We also characterized the cochlear function of *Tmc1*^{Bth/+} mice eight weeks after treatment. Mean ABR thresholds following Cas9-Tmc1-mut3-lipid injection remained lower than uninjected controls from 5.7–23 kHz, although the average improvement was lower than at four weeks post-treatment (Extended Data Fig. 4c, d), potentially owing to continued progressive hearing loss in the non-edited hair cells.

As a behavioural measure of hearing rescue, we assessed acoustic startle responses eight weeks after injection. In uninjected *Tmc1*^{Bth/+} mice, no startle response was detected following stimulation at 120 dB. By contrast, significant startle responses were detected in Cas9-Tmc1-mut3-lipid-injected *Tmc1*^{Bth/+} mice following stimulus at 110 and 120 dB (Fig. 3d and Extended Data Fig. 4e), demonstrating that hearing preservation upon treatment also preserves an acoustic behavioural reflex.

To evaluate the ability of each of the other *Tmc1*^{Bth}-targeting sgRNAs to mediate hearing rescue *in vivo*, we also injected Tmc1-mut1, Tmc1-mut2, and Tmc1-mut4 complexed with Cas9 into neonatal *Tmc1*^{Bth/+} cochleae, and observed varying degrees of enhanced cochlear function (Extended Data Fig. 7). Thus, while Tmc1-mut3 resulted in the most robust hearing preservation, other sgRNAs targeting the mutant *Bth* allele also partially preserved cochlear function.

To test whether RNP delivery of editing agents in adult mouse inner ears supports genome editing in hair cells, we injected Cas9-GFP sgRNA-lipid complexes into the cochleae of six-week-old Atoh1-GFP

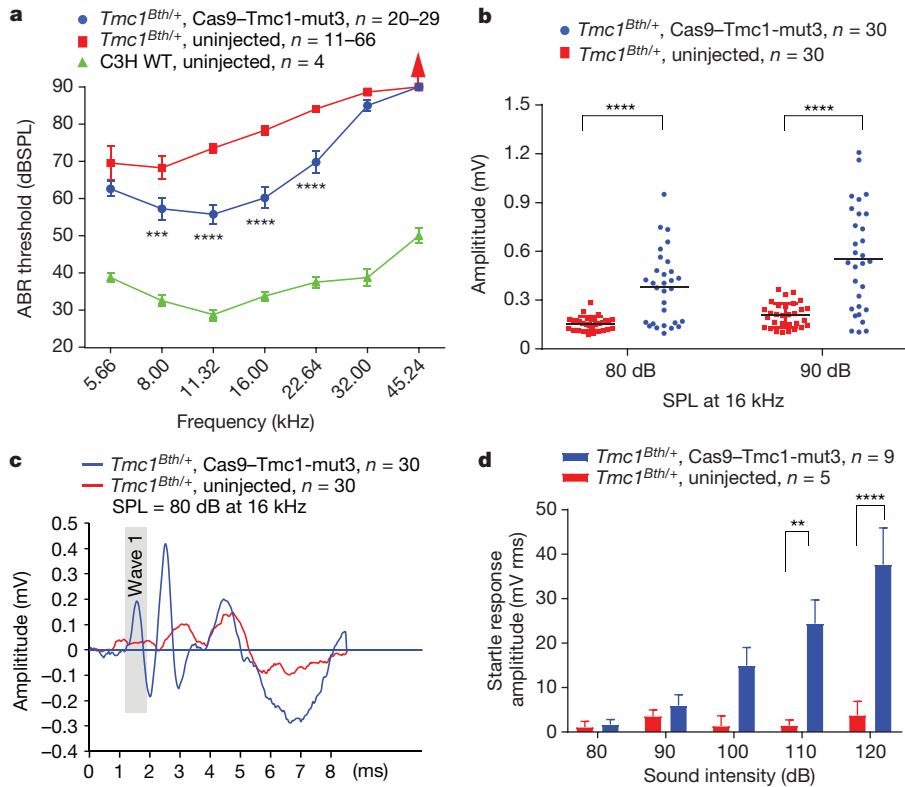


Figure 3 | Cas9-Tmc1-mut3-lipid injections reduce hearing loss in *Tmc1*^{Bth/+} mice. **a**, ABR thresholds in *Tmc1*^{Bth/+} ears injected with Cas9-Tmc1-mut3-lipid (blue), uninjected *Tmc1*^{Bth/+} ears (red), and wild-type C3H ears (green) after four weeks. **b**, Peak amplitudes of ABR wave 1 at 16 kHz in Cas9-Tmc1-mut3-lipid-injected ears (blue) compared with uninjected ears (red) after four weeks. Horizontal bars are mean values. **c**, Mean ABR waveforms in Cas9-Tmc1-mut3-lipid-injected ears (blue) and uninjected ears (red). **d**, Startle responses in Cas9-Tmc1-mut3-lipid-injected mice (blue) and in uninjected mice (red) eight weeks after treatment. Red arrow in **a** indicates no ABR response at the highest stimulus level tested (90 dB). Statistical tests were two-way ANOVA with Bonferroni correction for multiple comparisons: ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001. Values and error bars reflect mean ± s.e.m. Among the different frequencies assayed, the number of ears tested (*n*) varies within the range shown (Supplementary Table 2).

mice by canalostomy. Two weeks after injection, loss of GFP fluorescence in the apical turn suggested target gene disruption with $25 \pm 2.1\%$ efficiency (Extended Data Fig. 8), comparable to previous observations of 20% GFP editing in neonatal hair cells¹⁵. These results suggest that this approach may be applicable to dominant genetic deafness that manifests with late-onset hearing loss.

To confirm that *in vivo* treatment of *Tmc1*^{Bth/+} mice with Cas9-Tmc1-mut3 sgRNA disrupted the *Tmc1*^{Bth} allele, we sequenced DNA from cochlea tissue collected from injected *Tmc1*^{Bth/+} and untreated *Tmc1*^{Bth/+} mice. After injection on P1, tissues were removed on P5 and

separated into organ of Corti (containing hair cells), spiral ganglion, and spiral ligament samples (Extended Data Fig. 9a, b). We estimated the fraction of hair cells in dissected cochlear tissue to be only about 1.5% of the total cells used for DNA sequencing (Extended Data Fig. 9a, b). Nevertheless, we observed unambiguous indels at the *Tmc1*^{Bth} locus in cochlear tissue from treated mice (Fig. 4a). The organ of Corti samples contained, on average, *Tmc1* editing of 0.92% of total sequenced DNA, which corresponds to about 1.8% *Tmc1*^{Bth} allele disruption in the heterozygous mice (Fig. 4a). We also isolated samples of much smaller numbers of cells (up to a few dozen, mostly hair cells) from

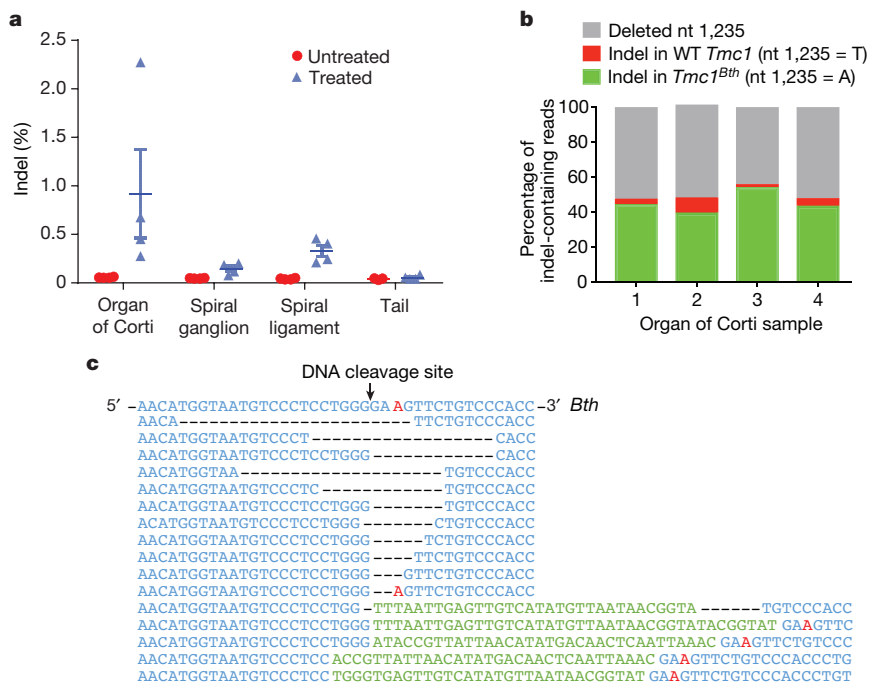


Figure 4 | Genome modification at *Tmc1* induced by lipid-mediated delivery of Cas9-Tmc1-mut3 RNP into *Tmc1*^{Bth/+} mice. **a**, *Tmc1* indel frequencies from tissue samples four days after injection of Cas9-Tmc1-mut3-lipid (blue) or from uninjected mice (red). Individual values (*n* = 4) are shown; horizontal lines and error bars reflect mean ± s.e.m. Note that *Tmc1*^{Bth} allele editing frequencies in these heterozygous mice are approximately double the observed indel frequencies. **b**, Analysis of indel-containing *Tmc1* sequencing reads from four injected organ of Corti samples in **a**. **c**, The most abundant 16 *Tmc1* sequences, grouped by similarity, from organ of Corti samples in **b**. The T1235A *Tmc1*^{Bth} mutation is shown in red.

treated mice. Decreasing the number of cells entering the genomic DNA amplification and sequencing process increased the observed editing percentage to as high as 10% *Tmc1*^{Bth} allele disruption, but also elevated background *Tmc1* indel rates of untreated mice to an average of $0.82 \pm 0.57\%$ and a maximum of 1.6%, probably reflecting increased noise from processing of minute quantities of genomic DNA. No indel frequencies above that of untreated controls at any of the above-identified off-target sites were observed in Cas9–*Tmc1*–mut3–lipid-treated tissues (Extended Data Fig. 9c). Together, these observations confirm that Cas9–*Tmc1*–mut3–lipid treatment *in vivo* edits the *Tmc1* locus with no detected editing at GUIDE-seq-identified off-target loci.

An analysis of indel-containing *Tmc1* sequencing reads from treated *Tmc1*^{Bth/+} mice allowed us to directly assess the allele specificity of Cas9–*Tmc1*–mut3 *in vivo*. Of 11,694 sequencing reads containing indels from four treated organ of Corti samples, 6,118 (52%) contained an intact nucleotide at *Tmc1* position 1,235. Of these, 5,736 (94%) contained modification of the mutant *Tmc1*^{Bth} allele, whereas only 382 (6%) contained modification of the wild-type *Tmc1* allele (Fig. 4b). Therefore, samples after treatment on average contained 15-fold higher modification of the *Tmc1*^{Bth} allele over the wild-type allele (Fig. 4b, c). These results demonstrate selective disruption of the *Tmc1*^{Bth} allele in *Tmc1*^{Bth/+} mice, consistent with observed hearing phenotypes, even though the *Tmc1*^{Bth} and wild-type *Tmc1* alleles differ only at a single base pair.

This work shows that cationic lipid-mediated Cas9–sgRNA complex delivery *in vivo* can achieve allele-specific gene disruption in a mouse model of a human genetic disease, resulting in amelioration of the disease phenotype. Our results suggest that this approach has potential for the treatment of autosomal-dominant hearing loss related to hair cell dysfunction, and provide a complementary strategy to other approaches that use antisense oligos (ASOs) or RNA interference^{6,25}. The genome editing strategy developed here may inform the future development of a DNA-free, virus-free, one-time treatment for certain genetic hearing loss disorders.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 4 April; accepted 24 November 2017.

Published online 20 December 2017.

- Angeli, S., Lin, X. & Liu, X. Z. Genetics of hearing and deafness. *Anat. Rec. (Hoboken)* **295**, 1812–1829 (2012).
- Marazita, M. L. *et al.* Genetic epidemiological studies of early-onset deafness in the U.S. school-age population. *Am. J. Med. Genet.* **46**, 486–491 (1993).
- Morton, C. C. & Nance, W. E. Newborn hearing screening—a silent revolution. *N. Engl. J. Med.* **354**, 2151–2164 (2006).
- Géléoc, G. S. & Holt, J. R. Sound strategies for hearing restoration. *Science* **344**, 1241062 (2014).
- Müller, U. & Barr-Gillespie, P. G. New treatment options for hearing loss. *Nat. Rev. Drug Discov.* **14**, 346–365 (2015).
- Shibata, S. B. *et al.* RNA interference prevents autosomal-dominant hearing loss. *Am. J. Hum. Genet.* **98**, 1101–1113 (2016).
- Pan, B. *et al.* Gene therapy restores auditory and vestibular function in a mouse model of Usher syndrome type 1c. *Nat. Biotechnol.* **35**, 264–272 (2017).
- Ahmed H., Shubina-Oleinik O. & Holt, J. R. Emerging gene therapies for genetic hearing loss. *J. Assoc. Res. Otolaryngol.* **18**, 649–670 (2017).
- Sacheli, R., Delacroix, L., Vandenackerveken, P., Nguyen, L. & Malgrange, B. Gene transfer in inner ear cells: a challenging race. *Gene Ther.* **20**, 237–247 (2013).

- Komor, A. C., Badran, A. H. & Liu, D. R. CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* **168**, 20–36 (2017).
- Long, C. *et al.* Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351**, 400–403 (2016).
- Yang, Y. *et al.* A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice. *Nat. Biotechnol.* **34**, 334–338 (2016).
- Yin, H. *et al.* Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components *in vivo*. *Nat. Biotechnol.* **34**, 328–333 (2016).
- Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J. S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**, 1012–1019 (2014).
- Zuris, J. A. *et al.* Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing *in vitro* and *in vivo*. *Nat. Biotechnol.* **33**, 73–80 (2015).
- Pan, B. *et al.* TMC1 and TMC2 are components of the mechanotransduction channel in hair cells of the mammalian inner ear. *Neuron* **79**, 504–515 (2013).
- Camp, G. V. & Smith, R. *Hereditary Hearing Loss Homepage* <http://hereditaryhearingloss.org> (2017).
- Zhao, Y. *et al.* A novel DFNA36 mutation in TMC1 orthologous to the Beethoven (Bth) mouse associated with autosomal dominant hearing loss in a Chinese family. *PLoS One* **9**, e97064 (2014).
- Kurima, K. *et al.* Dominant and recessive deafness caused by mutations of a novel gene, TMC1, required for cochlear hair-cell function. *Nat. Genet.* **30**, 277–284 (2002).
- Kawashima, Y., Kurima, K., Pan, B., Griffith, A. J. & Holt, J. R. Transmembrane channel-like (TMC) genes are required for auditory and vestibular mechanosensation. *PLoS Arch.* **467**, 85–94 (2015).
- Vreugde, S. *et al.* Beethoven, a mouse model for dominant, progressive hearing loss DFNA36. *Nat. Genet.* **30**, 257–258 (2002).
- Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
- Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015).
- Zhang, F. *CRISPR Design* <http://crispr.mit.edu> (2015).
- Lentz, J. J. *et al.* Rescue of hearing and vestibular function by antisense oligonucleotides in a mouse model of human deafness. *Nat. Med.* **19**, 345–350 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by DARPA HR0011-17-2-0049 (to D.R.L.), US NIH R01 EB022376 (to D.R.L.), R35 GM118062 (to D.R.L.), R01 DC006908 (to Z.-Y.C.), P30 DC05209 (to M.C.L.), R01 DC00138 (to M.C.L.), and R01 DC013521 (to J.R.H.). We are grateful for support from the David-Shulsky Foundation (to Z.-Y.C.), a Frederick and Ines Yeatts Hair Cell Regeneration grant (to Y.T., V.L., M.H., and Y.S.), the Bertarelli Foundation and the Jeff and Kimberly Barber Fund (to J.R.H.), the Broad Institute (to D.R.L. and Z.-Y.C.), and the HHMI (to D.R.L.). We thank H. Rees, S. Tsai, M. Packer, K. Zhao and D. Usanov for assistance.

Author Contributions X.G. and Y.T. designed the research, performed genome editing (X.G.) and hearing biology (Y.T.) experiments, analysed data, and wrote the manuscript. V.L., M.H., W.-H.Y., B.P., Y.-J.H., and H.W. designed experiments, performed hearing biology experiments and analysed data. D.B.P., M.C.L., and W.-J.K. designed hearing experiments and analysed data. Y.S. and S.Y. supported hearing biology experiments. J.H.H. analysed GUIDE-seq data. D.B.T. supported genome editing experiments. Y.L. and Q.X. designed and synthesized lipids. J.R.H., Z.-Y.C. and D.R.L. designed and supervised the research, and wrote the manuscript. All authors edited the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to D.R.L. (drliu@fas.harvard.edu) and Z.Y.C. (Zheng-Yi_Chen@meei.harvard.edu).

METHODS

Primary cell culture. Wild-type, *Tmc1^{Bth/+}* and *Tmc1^{Bth/Bth}* fibroblasts were obtained from P5 pups (see below). Mice were euthanized and cleaned with 70% ethanol. Underarm skin fragments (1–2 cm²) were excised and submerged in cold HBSS (ThermoFisher). Subcutaneous fat was removed by forceps. Skin fragments were cut into ~1-mm² pieces with a 25G 5/8" syringe (1180125058, Covidien). Tissues were digested with 0.5 mg/ml Liberase DL (Sigma 5401160001) at 37 °C for 1 h with occasional pipetting up and down to break cell clumps. Warm culture medium (1:1 DMEM:F12 medium (ThermoFisher) with 15% fetal bovine serum (FBS) (ThermoFisher) and 100 U/ml penicillin+streptomycin (ThermoFisher) was added to stop the enzyme digestion. The solution was filtered with a 70-µm cell strainer (Falcon) and centrifuged at 200g for 5 min. The pellet was resuspended in culture medium and transferred to a 25-ml culture flask, then incubated at 37 °C with 5% CO₂ and 3% O₂. Fibroblasts were cultured for about 2–3 days to reach ~90% confluence, then passaged in 100-ml flasks in DMEM plus GlutaMax (ThermoFisher) supplemented with 10% (v/v) FBS at 37 °C with 5% CO₂.

Delivery of proteins complexed with cationic lipids into mouse fibroblasts. Cultured fibroblast cells were plated in 24-well format (500 µl well volume) in Dulbecco's modified Eagle's medium plus GlutaMAX (DMEM, Life Technologies) with 10% FBS (no antibiotics) at a cell density sufficient to reach ~80% confluence at the time of usage. Purified sgRNA was incubated with Cas9 protein for 5 min before complexing with cationic lipid^{15,26}. Delivery of Cas9–sgRNA was performed by combining 100 nM RNP complex with 3 µl cationic lipid in 50 µl OPTIMEM medium (Life Technologies) according to the manufacturer's protocol for DNA plasmid transfection. The above mixture containing cationic lipid and RNP was then added to cells. All complexing steps were performed at room temperature. Cells were harvested and genomic DNA was extracted for sequencing ~96 h after treatment.

GUIDE-seq and data analysis. Mouse fibroblasts were transfected using 1,000 ng Cas9 plasmid (pCas9), 300 ng sgRNA plasmid (pTmc1-mut3 sgRNA), and 50 pmol GUIDE-seq double-stranded oligodeoxynucleotides (dsODN) using a LONZA 4D-Nucleofector. Transfection programs were optimized following the manufacturer's instructions (CA158 and CA189, P2 Primary Cell 4D-Nucleofector X Kit). pmaxGFP Control Vector (400 ng; LONZA) was added to the nucleofection solution to assess nucleofection efficiency in primary cells. The medium was replaced ~16 h after nucleofection and cells were collected for genomic DNA extraction after ~96 h. For GUIDE-seq off-target DNA cleavage analysis, pCas9, pTmc1-mut3 sgRNA, pmaxGFP, and dsODN were nucleofected into *Tmc1^{Bth/+}* heterozygous mouse primary fibroblasts. A sample nucleofected with dsODN only served as a negative control. About 400 ng genomic DNA for each sample was sheared acoustically using a Covaris m220 sonicator to an average length of 500 bp in 130 µl TE buffer. Each sample was sequenced on an Illumina Miseq following previously described protocols²³. Reads were consolidated first by their Illumina indexes and then by the 8-nt molecular index that defines a single pre-PCR template fragment. The consolidated reads were mapped to the mouse reference genome (GRCm38) using BWA-MEM. Off-target sites were identified by first mapping the start position of the amplified sequences using a 10-bp sliding window, then retrieving the reference sequence around the site. Given the size of some of the deletions, the number of base pairs used as the flanking sequence was increased to 100 bp. The retrieved sequences were aligned to the Cas9 target sequence using a Smith–Waterman local-alignment algorithm.

High-throughput DNA sequencing of genomic DNA samples. Treated cells or tissues were collected after four days and genomic DNA was isolated using the Agencourt DNAdvance Genomic DNA Isolation Kit (Beckman Coulter) according to the manufacturer's instructions. On-target and off-target genomic regions of interest were amplified by PCR with flanking HTS primer pairs (listed in Supplementary Sequences). PCR amplification was carried out with Phusion high-fidelity DNA polymerase (ThermoFisher) according to the manufacturer's instructions using ~100 ng genomic DNA as a template. PCR cycle numbers were chosen to ensure the reaction was stopped during the log-linear range of amplification. PCR products were purified using RapidTips (Diffinity Genomics). Purified DNA was amplified by PCR with primers containing sequencing adaptors. The products were purified by gel electrophoresis and quantified using the Quant-iTTM PicoGreen dsDNA Assay Kit (ThermoFisher) and KAPA Library Quantification Kit-Illumina (KAPA Biosystems). Samples were sequenced on an Illumina MiSeq as previously described²⁷.

Sequencing reads were demultiplexed using MiSeq Reporter (Illumina), and individual FASTQ files were analysed with a custom Matlab script (Supplementary Note). Each read was pairwise aligned to the appropriate reference sequence using the Smith–Waterman algorithm. Base calls with a Q-score below 31 were excluded from calculating editing frequencies. Sequencing reads were scanned for exact matches to two 10-bp sequences that flank both sides of a window in which indels

might occur. If no exact matches were located, the read was excluded from analysis. If the length of this indel window exactly matched the reference sequence, the read was classified as not containing an indel. If the indel window was one or more bases longer or shorter than the reference sequence, then the sequencing read was classified as an insertion or deletion, respectively.

General *in vivo* experiments. All *in vivo* experiments were carried out in accordance with NIH guidelines for the care and use of laboratory animals and were approved by the Massachusetts Eye & Ear Infirmary IACUC committee. Isogenic heterozygous *Tmc1^{Bth/+}* mice maintained on a C3HeB/FeJ (C3H) background were obtained as a gift from A. Griffith²¹, and inbred with wild-type C3H mice obtained from Jackson Laboratory. Crossbred homozygous C3H-*Tmc1^{Bth/Bth}* mice were caged with C3H mice to generate heterozygous *Tmc1^{Bth/+}* mice. All mice were genotyped by Transnetyx. For mechanotransduction experiments, two genotypes of *Tmc* mutant mice: *Tmc1^{Bth/Bth}Tmc2^{Δ/Δ}* and *Tmc1^{Δ/Δ}Tmc2^{Δ/Δ}*¹⁶, were bred to generate *Tmc1^{Bth/Δ}Tmc2^{Δ/Δ}* mice.

Microinjection into the inner ear of neonatal mice. A total of 106 *Tmc1^{Bth/+}* or C3H mice (P0–2) of either sex were used for injections. The mice were randomly assigned to the different experimental groups. The final 25% of the experiments were performed in a double-blinded manner. At least five mice were injected in each group. All surgical procedures were done in a clean, dedicated space. Instruments were thoroughly cleaned with 70% ethanol and autoclaved before surgery. Fresh Cas9 and sgRNA were mixed before injection at a final concentration of 25 µM. One microlitre Lipofectamine 2000 was mixed with 1 µl Cas9–sgRNA RNP and incubated for 20 min at room temperature. Mice were anaesthetized by hyperthermia on ice. Cochleostomy was performed by preauricular incision to expose the cochlear bulla. Anatomical landmarks included the stapedia artery and tympanic ring, which were identified before injection. Glass micropipettes (4878, WPI) were pulled with a micropipette puller (PP83, Narishige) to a final outer diameter of ~10 µm. Needles held by a Nanolitre 2000 micromanipulator (WPI) were used to manually deliver the Cas9–sgRNA–lipid complexes into the scala media, which allows access to inner ear cells. The injection sites were the base, middle, and apex–middle turn of the cochlea. The volume for each injection was 0.3 µl with a total volume of 0.9 µl per cochlea. The release rate was 69 nl/min, controlled by a MICRO4 microinjection controller (WPI).

Microinjection into adult inner ear by canalostomy. Three 6-week-old *Atoh1*–GFP mice²⁸ were injected with Cas9–GFP sgRNA–lipid complex, with the same concentration and volume for each component as used in injection into neonatal inner ear. Mice were anaesthetized by intraperitoneal injection of xylazine (10 mg/kg) and ketamine (100 mg/kg). The right post-auricular region was exposed by shaving and disinfected with 10% povidone iodine. For canalostomy, a 10-mm postauricular incision was made under the operating microscope, and the right pinna and the sternocleidomastoid muscle were extracted to expose the posterior semicircular canal (PSCC), located in the margin of the temporal bone. We used a Bonn microprobe (Fine Science Tools) to drill a small hole on the PSCC, then left it open for a few minutes until no obvious perilymph leakage was observed. The tip of the polyimide tube (inner diameter 0.0039 inches, outer diameter 0.0049 inches, Microtumen) was inserted into the PSCC towards the ampulla. The hole was sealed with tissue adhesive (3M Vetbond), and a lack of fluid leakage indicated the tightness of the sealing. The tubing was cut after injection, with approximately 5 mm of tubing left connected to the PSCC and sealed with tissue adhesive. The volume for each injection was 1 µl per cochlea. The release rate was 169 nl/min, controlled by MICRO4 microinjection controller (WPI). The skin was closed with 5-0 nylon suture (Ethicon Inc.). The total surgery time was approximately 20 min, including a 6-min injection period.

Acoustic testing. ABR and DPOAE were recorded as described previously²⁹ at 32 °C in a soundproof chamber. Mice of either sex were anaesthetized with xylazine (10 mg/kg, intraperitoneally (i.p.)) and ketamine (100 mg/kg, i.p.). Acoustic stimuli were delivered through a custom acoustic assembly consisting of two miniature dynamic electrostatic earphones (CDMG15008-03A, CUI) to generate primary tones and a miniature microphone (FG-23329-PO7, Knowles) to record ear-canal sound pressure near the eardrum. Custom LabVIEW software controlling National Instruments 24-bit soundcards (6052E) generated all ABR and DPOAE stimuli and recorded all responses.

For ABR measurements, needle electrodes were inserted at the vertex and ventral edge of the pinna, with a ground reference near the tail. ABR potentials were evoked with 5-ms tone pips (0.5-ms rise–fall with a cos² onset, delivered at 35/s). The response was amplified 10,000-fold, filtered (100 Hz–3 kHz passband), digitized, and averaged (1,024 responses) at each SPL. The sound level was raised in 5 dB steps from 30 dB below threshold up to 90 dB SPL at frequencies from 5.66–45.24 kHz (in half-octave steps). Following visual inspection of stacked waveforms, “threshold” was defined as the lowest sound pressure level (SPL) at which any wave could be detected. In general, thresholds were defined by three

independent observers. Wave 1 amplitude was defined as the difference between the average of the 1-ms pre-stimulus baseline and the wave 1 peak (P1), after additional high-pass filtering to remove low-frequency baseline shifts.

For DPOAE measurements, the cubic distortion product was measured in response to primaries f_1 and f_2 . The primary tones were set so that the frequency ratio (f_2/f_1) was 1.2 and so that the f_2 level was 10 dB below the f_1 level. For each f_2/f_1 primary pair, primaries were swept in 5-dB steps from 20 dB SPL to 80 dB SPL (for f_2). At each level, the amplitude of the DPOAE at $2f_1-f_2$ was extracted from the averaged spectra, along with the noise floor. Threshold was computed by interpolation as the f_2 level required to produce a DPOAE at 5 dB SPL.

Acoustic startle reflex. Mice were placed into a small, acoustically transparent cage resting atop a piezoelectric force plate in a sound attenuated booth. Acoustic stimuli and amplified force plate signals were encoded by a digital signal processor (Tucker-Davis Technologies, RX6) using LabView scripts (National Instruments). Mice were placed in silence for 2 min and 60 dB broadband white noise for 5 min to acclimate to the test environment before real measurements. Broadband white noise was presented at a background level of 60 dB SPL throughout the experiment and a 16-kHz tone was presented at randomized intervals from an overhead speaker (80 dB to 120 dB SPL, 20 ms duration with 0 ms onset and offset ramps). Ten repetitions were recorded for each of the intensities per test subject. Startle response amplitude was measured as the root mean square (RMS) voltage of the force plate signal shortly after sound presentation.

Immunohistochemistry and histology. Injected and non-injected cochleae were removed after animals were killed by CO₂ inhalation. Temporal bones were fixed in 4% paraformaldehyde at 4°C overnight, then decalcified in 120 mM EDTA for at least 1 week. The cochleae were dissected in pieces from the decalcified tissue for whole-mount immunofluorescence. Tissues were infiltrated with 0.3% Triton X-100 and blocked with 8% donkey serum for 1 h before applying the first antibody. Rabbit anti-MYO7A (1:500; #25-6790, Proteus BioSciences), chicken anti-GFP (1:750; ab13970, Abcam) and goat anti-SOX2 (1:350; sc-17320, Santa Cruz Biotechnology) were used at room temperature overnight. The second antibody was incubated for 1 h after three rinses with PBS rinses. All Alexafluor secondary antibodies were from Invitrogen: donkey anti-rabbit Alex488 (A21206) or Alex 594 (A21207), donkey anti goat Alex594 (A11058) or Alexa-488-phalloidin (A12379) and goat anti-chicken Alex488 (A-11039) were used at a 1:500 dilution. Specimens were mounted in ProLong Gold Antifade Mountant medium (P36930, Life Technologies). Confocal images were taken with a Leica TCS SP5 microscope using a 20× or 63× glycerin-immersion lens, with or without digital zoom. For IHC and OHC counting, we acquired z-stacks by maximum intensity projections of z-stacks for each segment by imageJ (NIH image), and composite images showing the whole cochlea were constructed in Adobe Photoshop CS3 to show the whole turn of cochlea. A frequency map was constructed for each case by measuring the spiral extent of all the dissected cochlear pieces and converting cochlear location to frequency using a plug-in of ImageJ (<https://www.masseyeandear.org/research/otolaryngology/investigators/laboratories/eaton-peabody-laboratories/epl-histology-resources/imagej-plugin-for-cochlear-frequency-mapping-in-whole-mounts>). MYO7A-positive IHCs and OHC were counted in the cochlear regions that respond to different sound frequencies, and any segments containing dissection-related damage were omitted from further analysis.

Hair cell transduction current recording. Wild-type or $Tmc1^{Bth/\Delta}Tmc2^{\Delta/\Delta}$ littermates were injected with 0.9 μ l Cas9–Tmc1–mut3–Lipofectamine 2000 or Cas9–GFP sgRNA–Lipofectamine 2000. Wild-type C57BL/6 mice were injected with 0.9 μ l Cas9–Tmc1–wt3–Lipofectamine 2000 at P0–P1 via cochleostomy. Cochleae were removed at P5–P6 and cultured in MEM(1×) + GlutaMAX-I medium with 1% FBS at 37°C, 5% CO₂ for up to 15 days. For recording, the organs of Corti were bathed in standard artificial perilymph containing 137 mM NaCl, 0.7 mM NaH₂PO₄, 5.8 mM KCl, 1.3 mM CaCl₂, 0.9 mM MgCl₂, 10 mM HEPES, and 5.6 mM D-glucose. Vitamins (1:50) and amino acids (1:100) were added to the solution from concentrates (Invitrogen, ThermoFisher Scientific), and NaOH was used to adjust the final pH to 7.4 (~310 mOsm/kg). Recording pipettes (2–4 M Ω)

were pulled from R6 capillary glass (King Precision Glass) and filled with intracellular solution containing 135 mM CsCl, 5 mM HEPES, 5 mM EGTA, 2.5 mM MgCl₂, 2.5 mM Na₂-ATP, and 0.1 mM CaCl₂; CsOH was used to adjust the final pH to 7.4 (~285 mOsm/kg). Whole-cell, tight-seal, voltage-clamp recordings were conducted at –84 mV at room temperature (22–24°C) with an Axopatch 200B amplifier (Molecular Devices). Hair bundles were deflected with a stiff glass probe fabricated from capillary glass with a fire polisher (MF-200, World Precision Instruments) to create a rounded probe tip of ~3–5 μ m in diameter. Probes were mounted on a PICMA Chip piezo actuator (Physik Instrument) and driven by an LVPZT amplifier (E-500.00, Physik Instrumente). Sensory-transduction currents were recorded from uninjected and Cas9–sgRNA-treated hair cells. The data were filtered at 10 kHz with a low-pass Bessel filter and digitized at >20 kHz with a 16-bit acquisition board (Digidata 1440A, Molecular Devices) and pClamp 10 software (Molecular Devices).

Inner ear tissue dissection for HTS. $Tmc1^{Bth/+}$ mice were injected with Cas9–sgRNA at P1 as described above. All dissection instruments were thoroughly cleaned with 70% ethanol and DNase Free (D6002, ARgos), then autoclaved before dissection. Mice were euthanized at P5. Temporal bones were removed and immersed in clean PBS pH 7.4 (10010001, ThermoFisher) individually. Different forceps were used for each ear. The organ of Corti, spiral ganglion, and spiral ligament from the injected and non-injected ear, and tail tissue were all removed under microscope from each mouse.

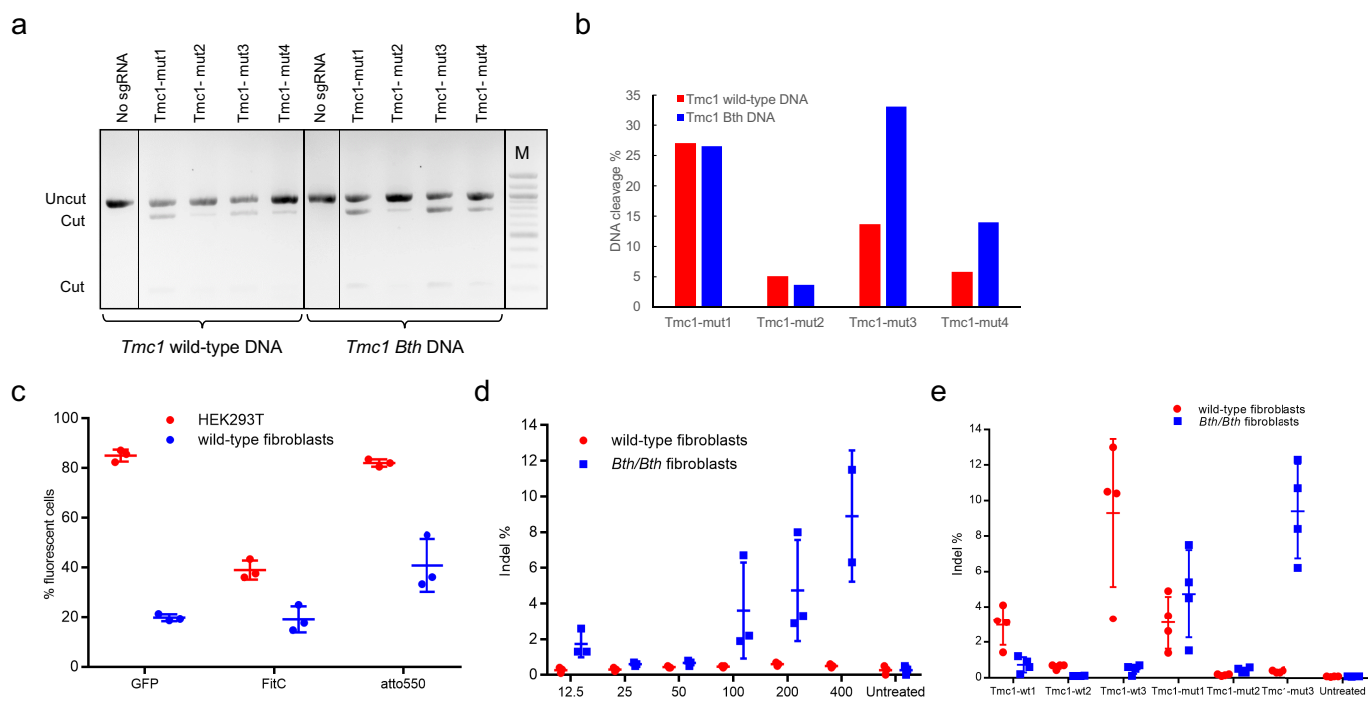
Hair cells isolation for HTS. $Tmc1^{Bth/+}$ mice were injected with Cas9–Tmc1–mut3–Lipofectamine 2000 at P1 and euthanized at P5. Cochleae were dissected and immersed in 1 μ M FM 1-43FX (PA1-915, ThermoFisher) dissolved in HBSS (ThermoFisher) for 10 s at room temperature in the dark. Cochleae were rinsed three times with HBSS and placed in 100 μ l Cell Recovery Solution (354253, Discovery Labware) for 10 min at 37°C, then transferred to 100 μ l TrypleE Express Enzyme (12604013, ThermoFisher). Sensory epithelia were extracted by forceps. After incubation for 10 min at 37°C, the tissues were pipetted up and down 30 times. FM 1-43-positive cells were isolated using a 1- μ l pipette under a microscope (Axiovert 200M, Carl Zeiss), then subjected to whole-genome amplification by MALBAC Single Cell WGA Kit (YK001A, Yikon Genomics).

Statistical analysis. Statistical analyses were performed by two-way ANOVA with Bonferroni corrections for multiple comparisons for ABRs, DPOAEs, and acoustic startle response; and by Student's *t*-test for hair cell transduction currents using Prism 6.0 (GraphPad). No statistical methods were used to predetermine sample size. A total of 106 $Tmc1^{Bth/+}$ or C3H mice (P0–2) of either sex were used for injections. The mice were randomly assigned to the different experimental groups. The final 25% of the experiments were performed in a double-blinded manner.

Code availability. Labview software for cochlear function testing is available here: <http://www.masseyeandear.org/research/otolaryngology/investigators/laboratories/eaton-peabody-laboratories/epl-engineering-resources>. Matlab scripts used to quantify the acoustic startle response are available from the corresponding authors on request. Indel identification scripts are provided in the Supplementary Information.

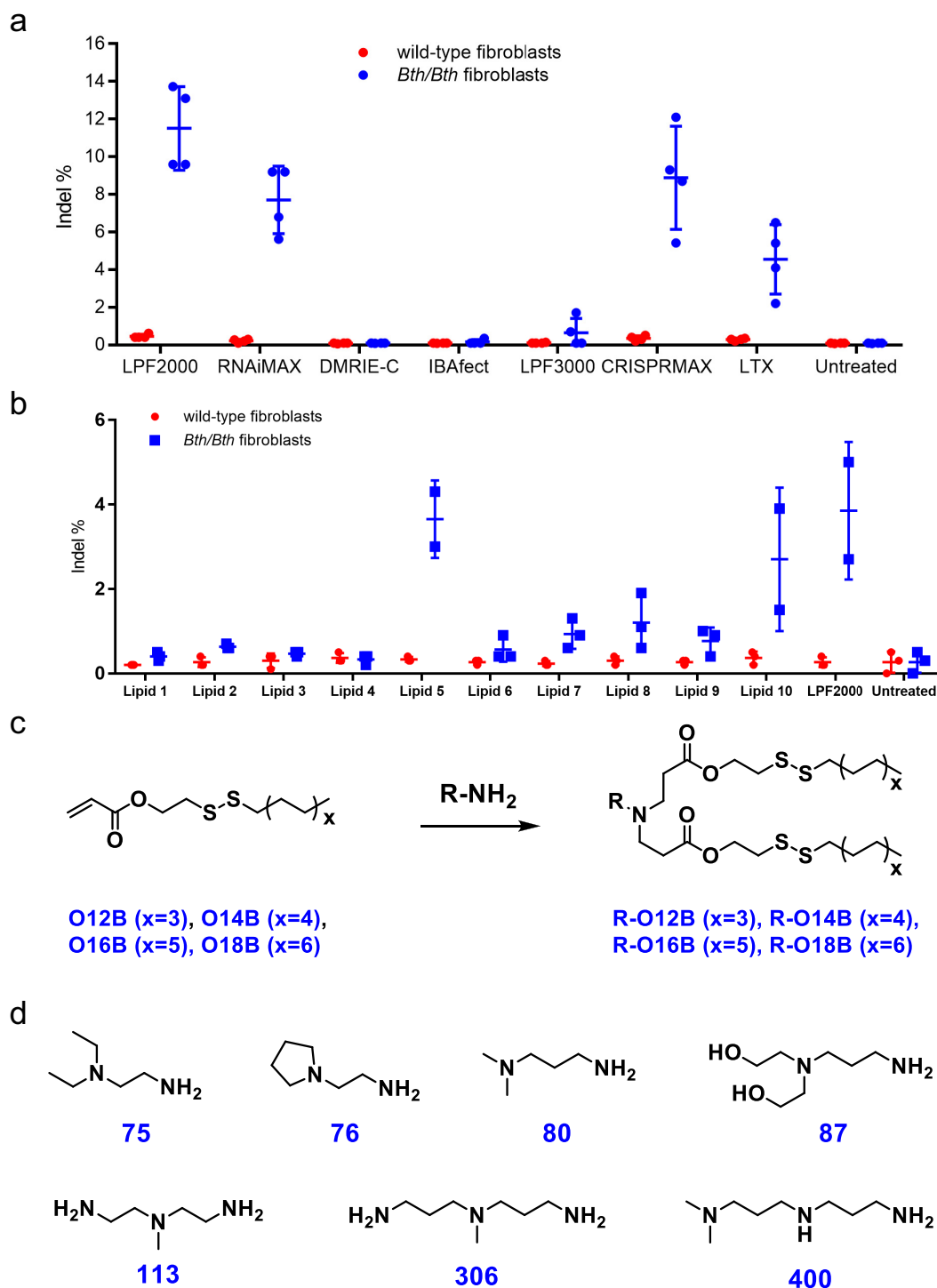
Data availability. High-throughput sequencing data have been deposited in the NCBI Sequence Read Archive database under accession code SRP103108. All other data are available from the corresponding authors on reasonable request.

26. Wang, M. *et al.* Efficient delivery of genome-editing proteins using bioreducible lipid nanoparticles. *Proc. Natl Acad. Sci. USA* **113**, 2868–2873 (2016).
27. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31**, 839–843 (2013).
28. Lumpkin, E. A. *et al.* Math1-driven GFP expression in the developing nervous system of transgenic mice. *Gene Expr. Patterns* **3**, 389–395 (2003).
29. Huang, M., Kantardzhieva, A., Scheffer, D., Liberman, M. C. & Chen, Z. Y. Hair cell overexpression of Islet1 reduces age-related and noise-induced hearing loss. *J. Neurosci.* **33**, 15086–15094 (2013).



Extended Data Figure 1 | Allele-selective editing of wild-type or *Bth* mutant *Tmc1* in cleavage assays *in vitro* and by lipid-mediated delivery into primary fibroblasts. a, *In vitro* Cas9-sgRNA-mediated *Tmc1* DNA cleavage. We incubated 100 nM of a 995-bp DNA fragment containing wild-type *Tmc1* (lanes 1–5) or *Tmc1*^{*Bth*} (lanes 6–10) with 300 nM of each of the four Cas9-sgRNAs shown for 15 min at 37 °C. Expected cleavage products are 774–778 bp and 217–221 bp. M, 100-bp ladder; the lower two heavy bands are 500 and 1,000 bp. **b**, Quantification of DNA cleavage in **a** by densitometry using imageJ. **c**, Comparison of transfection efficiency in HEK293T cells and wild-type primary fibroblasts. Fifty nanograms GFP plasmid, 10 nM Cas9-FitC-Tmc1-mut3 sgRNA RNP, or 10 nM Cas9-CrRNA-Tmc1-mut3-atto-550-TracrRNA RNP were delivered into HEK293T cells or wild-type primary fibroblasts using 3 μ l Lipofectamine 2000. For samples with GFP plasmid, the fraction of GFP-positive cells was measured by flow cytometry 24 h after delivery. For samples with Cas9-FitC-Tmc1-mut3 RNP or Cas9-CrRNA-Tmc1-mut3-atto-550-TracrRNA RNP, medium was removed 6 h after delivery. The cells were trypsinized,

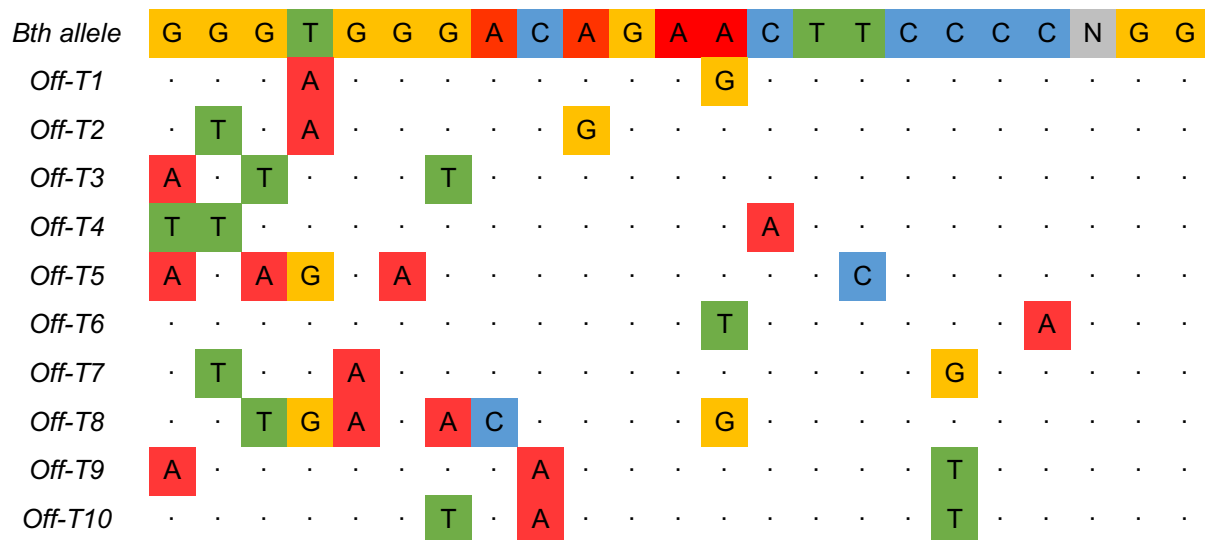
washed three times with 500 μ l PBS containing 20 U ml⁻¹ heparin, and subjected to flow cytometry. **d**, Wild-type or *Bth* mutant *Tmc1* allele editing in primary fibroblasts derived from wild-type or *Tmc1*^{*Bth/Bth*} mice as a function of the dose of Cas9-Tmc1-mut3-lipid complex. Cas9-Tmc1-mut3 (12.5, 25, 50, 100, 200, or 400 nM) was delivered into the primary fibroblasts using Lipofectamine 2000 in DMEM-FBS. **e**, Lipid-mediated delivery of Cas9-sgRNA complexes into primary fibroblasts derived from wild-type or *Tmc1*^{*Bth/Bth*} mice. We delivered 100 nM of purified Cas9 protein and each wild-type *Tmc1*-targeting sgRNA (Tmc1-wt1, Tmc1-wt2, or Tmc1-wt3) or *Tmc1*^{*Bth*} mutant-targeting sgRNA (Tmc1-mut1, Tmc1-mut2, or Tmc1-mut3) into wild-type fibroblasts (red) and *Tmc1*^{*Bth/Bth*} fibroblasts (blue) using Lipofectamine 2000 in DMEM-FBS. Primary fibroblast cells were harvested 96 h after treatment. Genomic DNA was extracted and indels were detected by HTS. Individual values ($n = 3-4$) are shown; horizontal lines and error bars represent mean \pm s.d. of biological replicates.



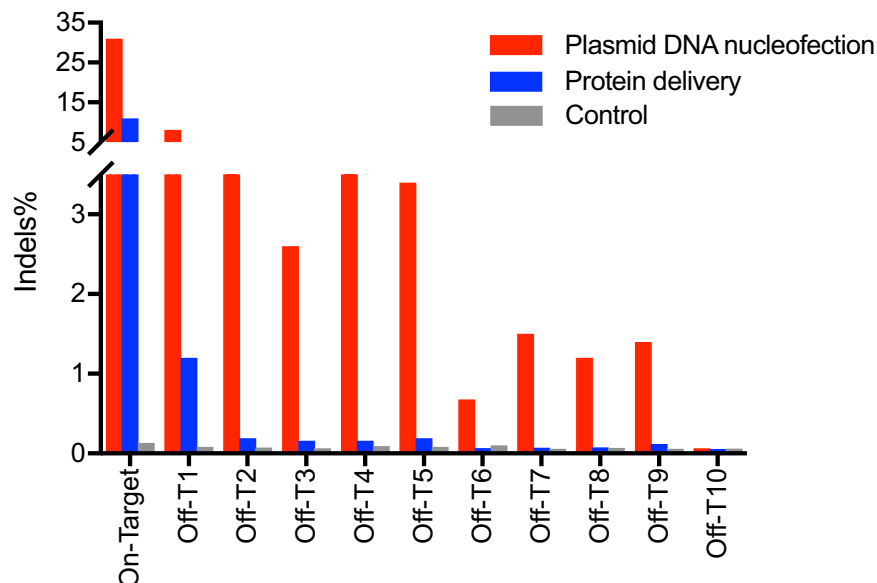
Extended Data Figure 2 | Delivery of Cas9–Tmc1–mut3 sgRNA complexes into primary fibroblasts derived from wild-type or homozygous *Tmc1*^{Bth/Bth} mice. **a**, Using seven commercially available lipids: LPF2000 (Lipofectamine 2000); RNAiMAX (Lipofectamine RNAiMAX); LPF3000 (Lipofectamine 3000); CRISPRMAX (Lipofectamine CRISPRMAX); LTX (Lipofectamine LTX). **b**, Using ten biodegradable, bio-reducible lipids: Lipid 1 (75-O14B); Lipid 2 (76-O14B); Lipid 3 (80-O18B); Lipid 4 (87-O16B); Lipid 5 (113-O18B); Lipid 6 (306-O12B); Lipid 7 (306-O16B); Lipid 8 (306-O18B);

Lipid 9 (400-O12B); Lipid 10 (400-O16B). We delivered 100 nM purified Cas9–Tmc1–mut3 RNP using 3 μ l of the cationic lipid shown in DMEM-FBS. Fibroblast cells were collected 96 h after treatment, genomic DNA was extracted, and indels were detected by HTS. **c**, Synthetic route and chemical structure of lipids. **d**, Commercially available amine head groups used in lipid synthesis. Lipids were synthesized as previously described²⁶. Individual values ($n = 2-4$) are shown; horizontal lines and error bars represent mean \pm s.d. of three or more biological replicates.

a

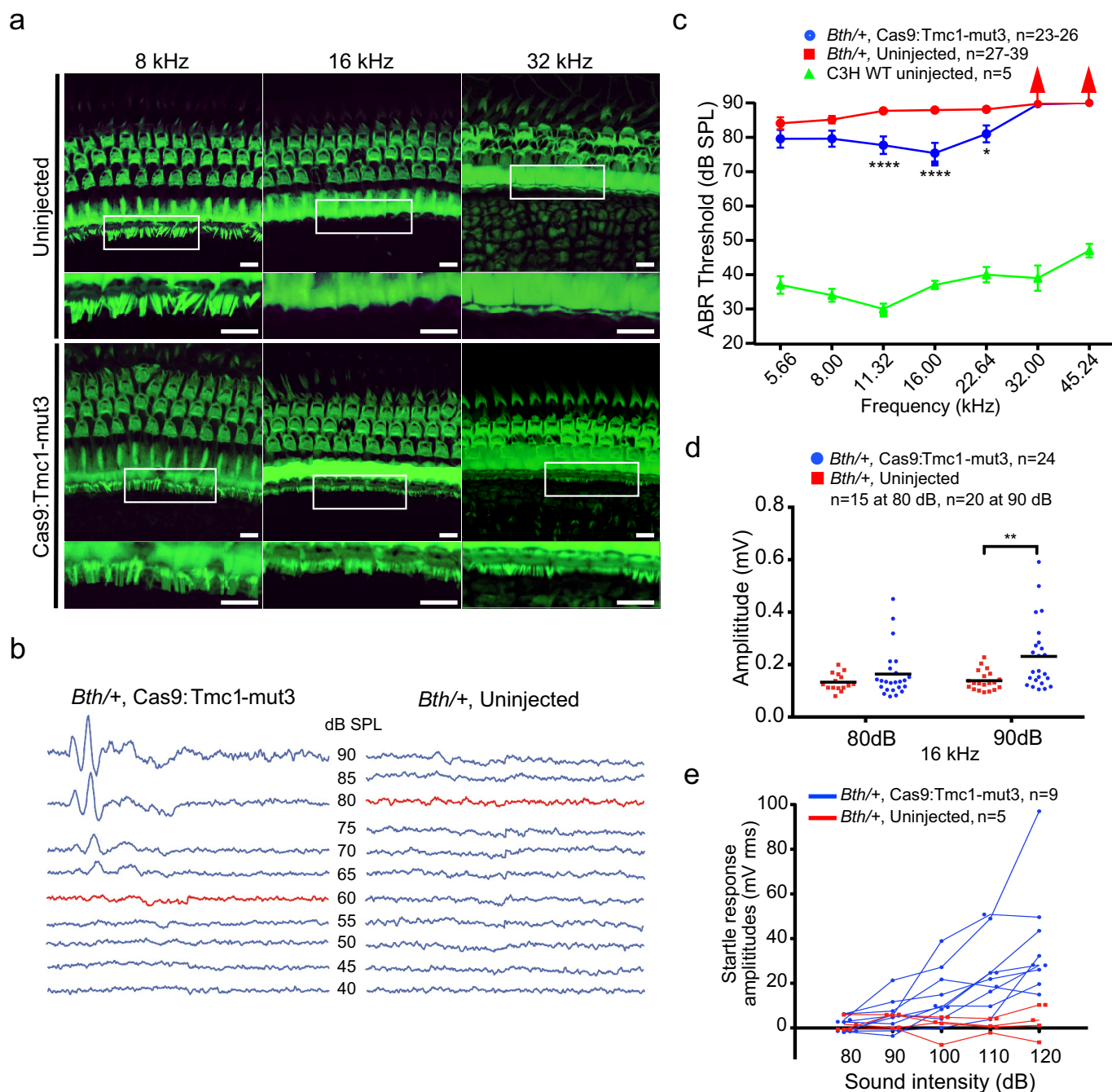


b



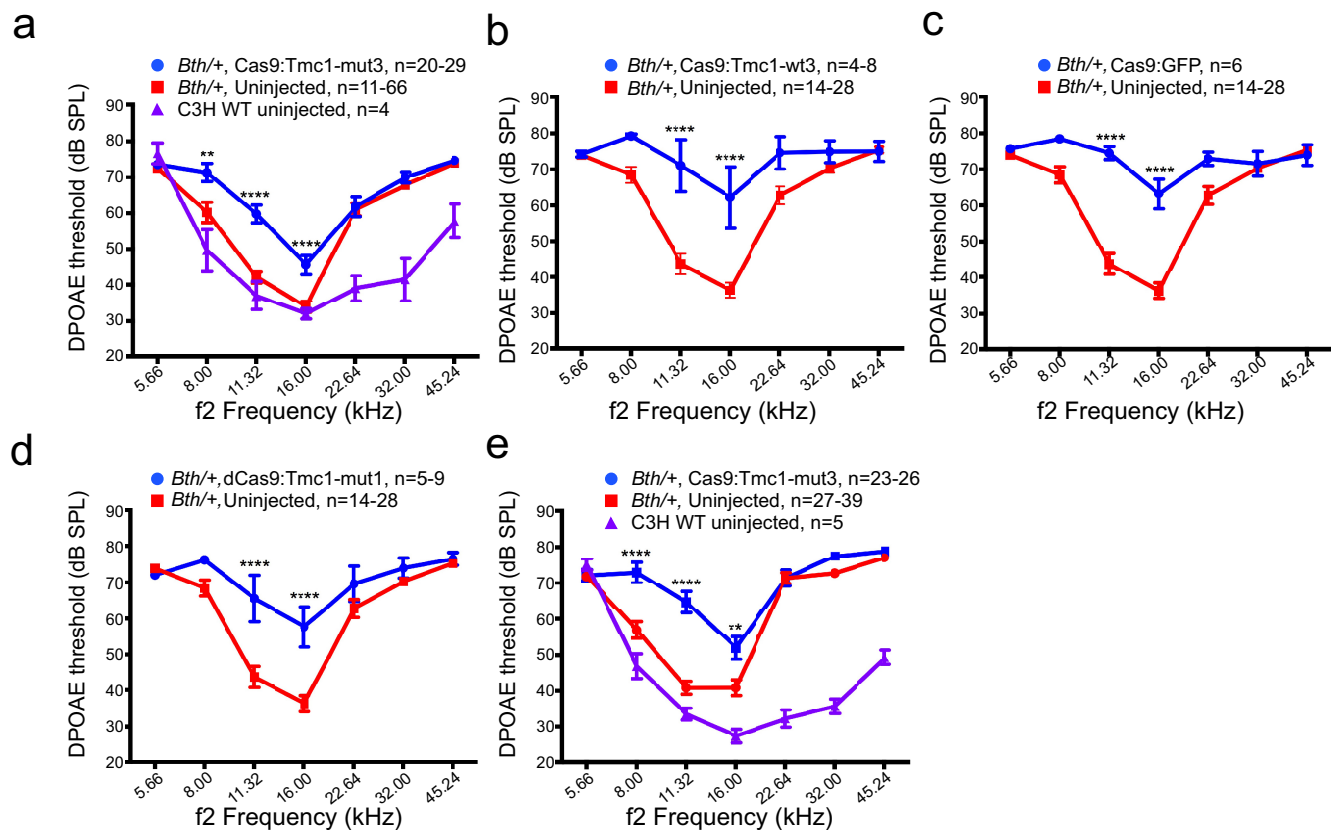
Extended Data Figure 3 | Off-target sites identified by GUIDE-seq after nucleofection of DNA plasmids encoding Cas9 and Tmc1-mut3 sgRNA into primary fibroblasts from *Tmc1^{Bth/+}* mice. a, One thousand nanograms of Cas9 plasmid, 300 ng Tmc1-mut3 sgRNA plasmid, 400 ng pmaxGFP plasmid, and 50 pmol double-stranded oligodeoxynucleotides (dsODN) were nucleofected into *Tmc1^{Bth/+}* fibroblasts using a LONZA 4D-Nucleofector. Genomic DNA was extracted 96 h after nucleofection and subjected to GUIDE-seq as previously described²³. *Off-T1* to *Off-T10* are ten off-target sites detected by GUIDE-seq. Mismatches compared to the on-target site are shown and highlighted in colour. The *Tmc1^{Bth}*

allele targeted by sgRNA Tmc1-mut3 is shown in the top row. **b**, Indel frequency at the *Tmc1* locus and at each of the off-target loci in Cas9–Tmc1-mut3-treated *Tmc1^{Bth/Bth}* primary fibroblasts following plasmid DNA nucleofection or following RNP delivery. For RNP delivery, 100 nM Cas9–Tmc1-mut3 RNP was delivered to the *Tmc1^{Bth/Bth}* fibroblasts using 3 μ l Lipofectamine 2000. Indels were detected by HTS at the *Tmc1* on-target site and at each off-target site. Red, samples nucleofected with DNA plasmids encoding Cas9 and Tmc1-mut3 sgRNA; blue, samples treated with Cas9–Tmc1-mut3 RNPs; grey, control samples nucleofected with unrelated dsDNA only.



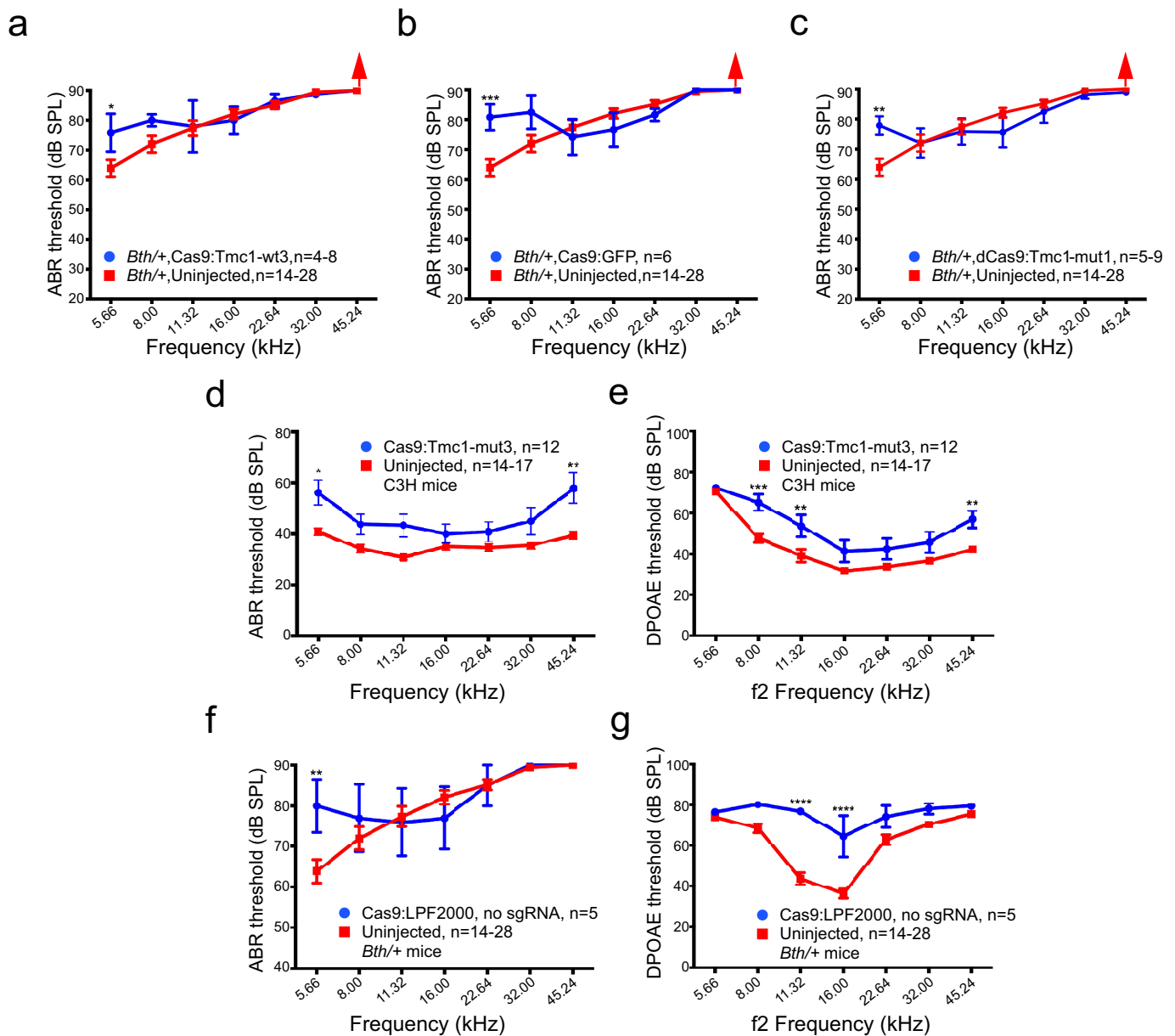
Extended Data Figure 4 | Cas9-Tmc1-mut3-lipid injection reduces hearing loss, improves acoustic startle response, and preserves stereocilia in *Tmc1*^{Bth/+} mice. **a**, Phalloidin labelling showed the preservation of stereocilia of IHCs in an ear eight weeks after injection with Cas9-Tmc1-mut3 sgRNA at three frequency locations indicated, whereas the uninjected contralateral inner ear of the same mouse showed severe degeneration of stereocilia at locations corresponding to 16 and 32 kHz. The boxes indicate the stereocilia, which are shown at the bottom of each image at higher magnification. Scale bars, 10 μ m. Similar results were observed in other injected ears that were immunolabelled ($n = 5$). **b**, Representative ABR waveforms showing reduced threshold (red traces) at 16 kHz in a Cas9-Tmc1-mut3-lipid-injected *Tmc1*^{Bth/+} ear (left) compared to the uninjected contralateral ear (right) of the same mouse after four weeks. **c**, Eight weeks after Cas9-Tmc1-mut3 injection into *Tmc1*^{Bth/+} ears (blue), mean ABR thresholds were significantly reduced at

three frequencies. Uninjected *Tmc1*^{Bth/+} ears (red) showed ABR thresholds > 85 dB at all frequencies after eight weeks. ABR thresholds from wild-type C3H mice are shown in green. **d**, ABR wave 1 amplitudes following 90 dB SPL stimulation at 16 kHz were greater in injected *Tmc1*^{Bth/+} ears than in uninjected ears eight weeks after treatment. Individual values ($n = 15$ or 20 for uninjected, and 24 for injected) are shown; horizontal bars represent mean values. **e**, Startle responses at 16 kHz in individual Cas9-Tmc1-mut3 sgRNA-injected mice (blue) were significantly stronger ($P < 0.001$) than in uninjected mice (red) eight weeks after treatment. Among the different frequencies assayed, the number of ears tested (n) varies within the range shown (Supplementary Table 2). Statistical analyses of ABR thresholds, amplitudes, and startle responses were performed by two-way ANOVA with Bonferroni correction for multiple comparisons: * $P < 0.05$, ** $P < 0.01$, **** $P < 0.0001$. Values and error bars reflect mean \pm s.e.m.



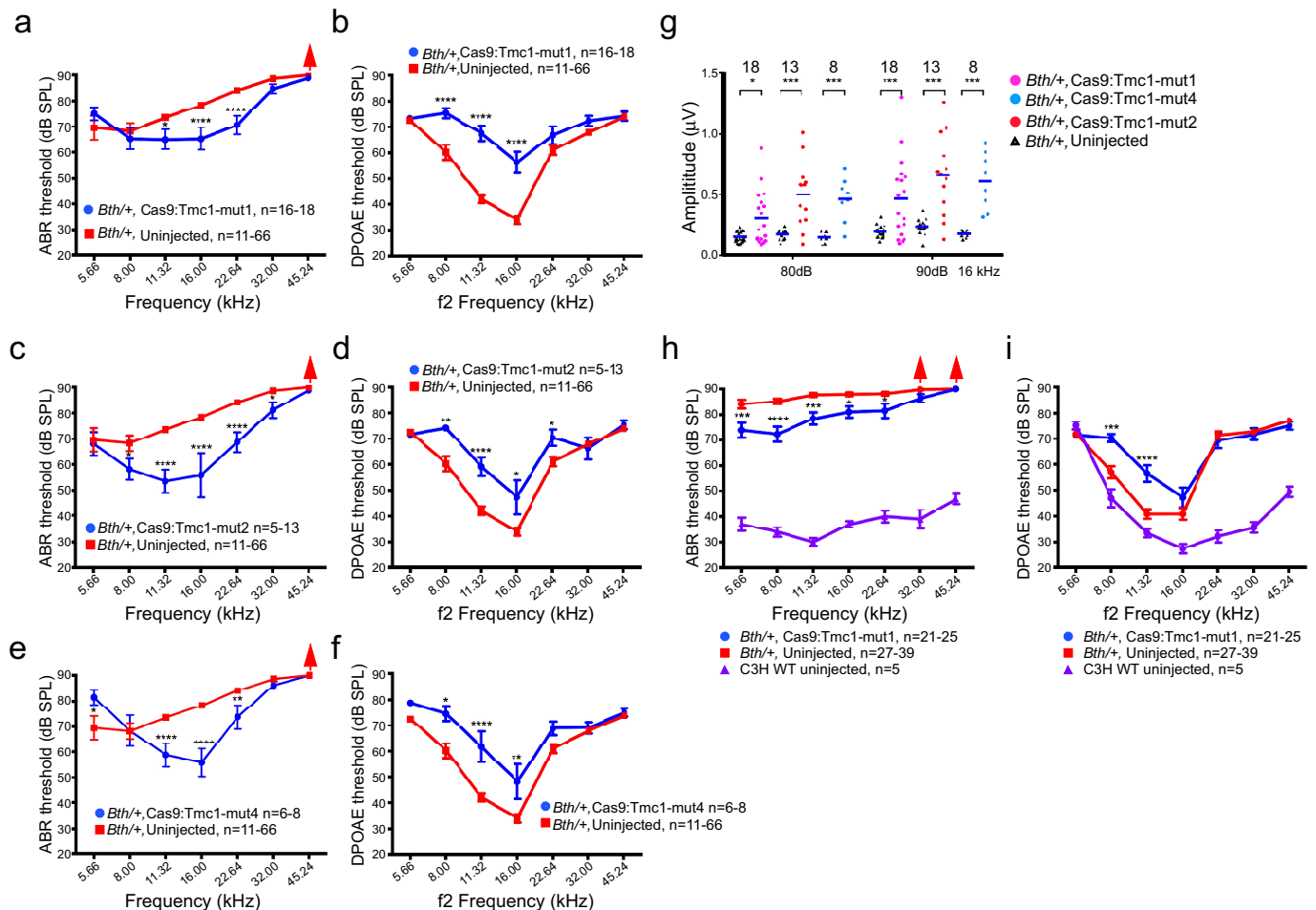
Extended Data Figure 5 | Effect of *in vivo* injection of Cas9-sgRNA-lipid complexes on DPOAE thresholds. **a–d**, DPOAE thresholds four weeks after injection were elevated compared with uninjected ears at three frequencies following treatment with Cas9-Tmc1-mut3 sgRNA (**a**), and were elevated at two frequencies following treatment with Cas9-Tmc1-wt3 sgRNA (**b**), Cas9-GFP sgRNA (**c**), or dCas9-Tmc1-mut1 sgRNA (**d**). **e**, Eight weeks after Cas9-Tmc1-mut3 sgRNA injection, DPOAE thresholds were elevated at three frequencies in the injected group. Mean DPOAE thresholds of untreated wild-type (WT) C3H mice at four weeks (**a**) or eight weeks (**e**) of age are also shown in purple. Statistical analysis of DPOAE thresholds was performed by two-way ANOVA with Bonferroni

correction for multiple comparisons: $**P < 0.01$, $****P < 0.0001$. Values and error bars reflect mean \pm s.e.m. Among the different frequencies assayed, the number of ears tested (n) varies within the range shown (Supplementary Table 2). The elevation of DPOAE thresholds despite enhanced hair cell survival (Fig. 2d, g) suggests that the surviving OHCs may not be fully functional. IHCs can respond to sound and excite auditory nerve fibres in the absence of OHC amplification, although at higher SPLs. Thus, an improvement in ABR thresholds and suprathreshold amplitudes can occur without concomitant DPOAE enhancement if the functional improvements are restricted to the surviving IHCs.



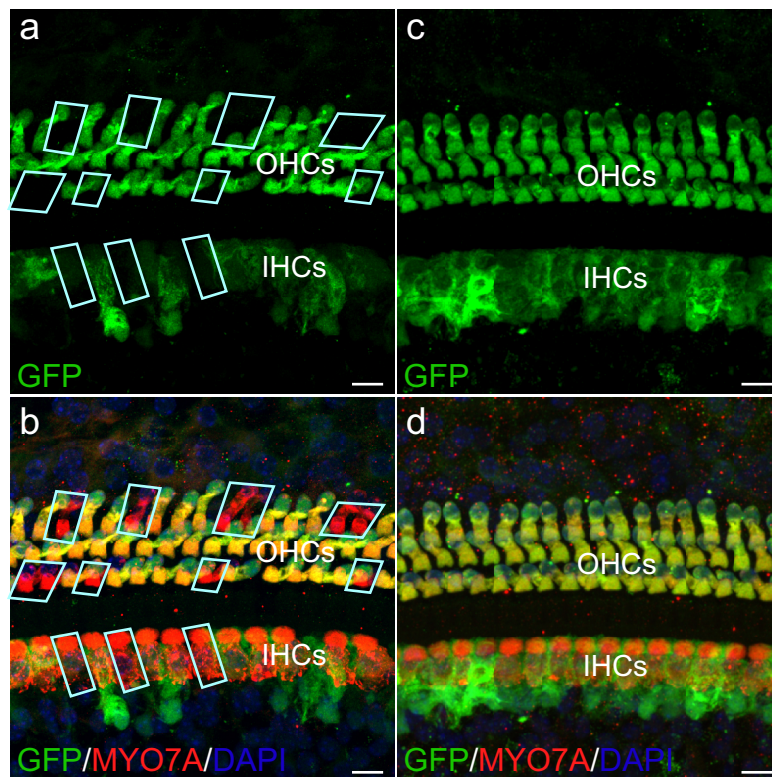
Extended Data Figure 6 | Hearing rescue is dependent on the *Tmc1*^{Bth} target specificity of the sgRNA, Cas9 nuclease activity, the presence of the *Tmc1*^{Bth} mutation, and the presence of the sgRNA. **a**, In *Tmc1*^{Bth/+} ears injected with Cas9-Tmc1-wt3-lipid, which targets the wild-type *Tmc1* allele instead of the mutant *Tmc1*^{Bth} allele, ABR thresholds (blue) were comparable to or higher than those of uninjected controls (red) after four weeks. **b**, *Tmc1*^{Bth/+} ears injected with Cas9-GFP sgRNA-lipid (blue) did not show improved ABR thresholds four weeks after treatment. **c**, *Tmc1*^{Bth/+} ears injected with catalytically inactive dCas9-Tmc1-mut1-lipid did not show improved ABR thresholds four weeks after treatment. **d**, ABR thresholds of wild-type C3H mice injected with Cas9-Tmc1-mut3-lipid showed similar patterns to the uninjected control inner

ears at four weeks, except at 5.66 and 45.24 kHz where ABR thresholds were elevated. **e**, Elevated DPOAE thresholds at three frequencies were observed after the treatment in **d**. **f**, Injection of Cas9-Lipofectamine 2000 (LPF2000) without sgRNA in *Tmc1*^{Bth/+} mice did not improve ABR thresholds after four weeks. **g**, Elevated DPOAE thresholds at 11 and 16 kHz were observed after the treatment in **f**. Statistical analysis of ABR and DPOAE thresholds was performed by two-way ANOVA with Bonferroni correction for multiple comparisons: **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001. Values and error bars reflect mean ± s.e.m. Among the different frequencies assayed, the number of ears tested (*n*) varies within the range shown (Supplementary Table 2).



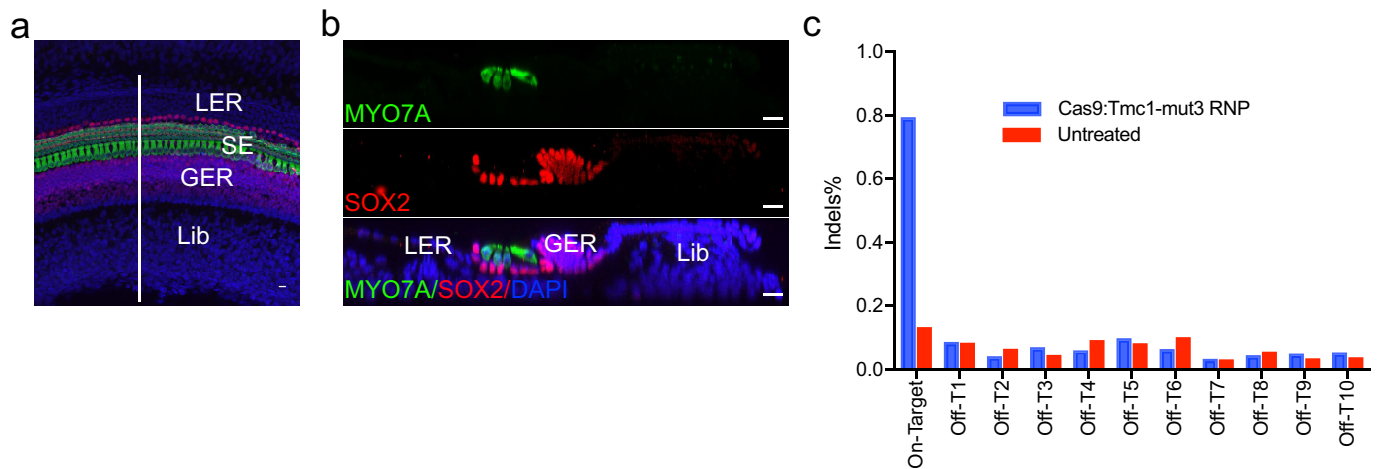
Extended Data Figure 7 | Hearing preservation following treatment with additional Tmc1-mut sgRNAs other than Tmc1-mut3. **a**, Mean ABR thresholds were significantly reduced at three frequencies in ears injected with Cas9–Tmc1-mut1–lipid compared to uninjected *Tmc1^{Bth/+}* ears after four weeks. **b**, DPOAE thresholds were elevated in the same group of inner ears after Cas9–Tmc1-mut1 injection as in **a** after four weeks. **c**, Mean ABR thresholds were significantly reduced at five frequencies in ears injected with Cas9–Tmc1-mut2–lipid compared to uninjected *Tmc1^{Bth/+}* ears after four weeks. **d**, DPOAE thresholds were elevated in the same group of inner ears after Cas9–Tmc1-mut2 injection as in **c** after four weeks. **e**, Mean ABR thresholds were significantly reduced at three frequencies in ears injected with Cas9–Tmc1-mut4–lipid compared to uninjected *Tmc1^{Bth/+}* ears after four weeks. **f**, DPOAE thresholds were elevated in the same group of inner ears after Cas9–Tmc1-mut4–lipid injection as in **e** after four weeks. **g**, Significantly stronger wave 1 amplitudes were detected in ears injected with each of the Cas9–Tmc1-mut–lipid complexes shown at 16 kHz (80 and 90 dB SPL). Individual

values ($n = 8, 13$, or 18) are shown; horizontal bars represent mean values. **h**, Eight weeks after Cas9–Tmc1-mut1–lipid injection into *Tmc1^{Bth/+}* ears, mean ABR thresholds were significantly reduced at five frequencies compared to uninjected *Tmc1^{Bth/+}* ears, which showed ABR thresholds >80 dB at all frequencies after eight weeks. Mean ABR thresholds of untreated wild-type (WT) C3H mice of eight weeks of age are shown in purple. Red arrows indicate no ABR response at the highest SPL level of 90 dB. **i**, DPOAE thresholds were significantly elevated at two frequencies (8 and 11 kHz) in the same group of inner ears after Cas9–Tmc1-mut1 injection as in **h** after eight weeks. Mean DPOAE thresholds of untreated wild-type C3H mice of eight weeks of age are shown in purple. Statistical analysis of ABR and DPOAE thresholds and wave 1 amplitudes was performed by two-way ANOVA with Bonferroni correction for multiple comparisons: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$. Values and error bars reflect mean \pm s.e.m. Among the different frequencies assayed, the number of ears tested (n) varies within the range shown (Supplementary Table 2).



Extended Data Figure 8 | RNP delivery of Cas9-sgRNA-lipid complexes results in genome editing in adult hair cells. Six-week-old adult Atoh1-GFP cochlea were injected with 1 μ l 25 μ M Cas9-GFP sgRNA-lipid complex by canalostomy, with the cochlea removed two weeks after injection. **a**, Genome editing was detected by the loss of GFP (green, with GFP absence noted using cyan shapes) in inner hair cells (IHCs)

and outer hair cells (OHCs). **b**, Hair cells were labelled with the hair cell marker MYO7A (red) in the apex turn of cochlea. **c**, **d**, In uninjected contralateral Atoh1-GFP cochlea, all hair cells were GFP-positive. Scale bars, 10 μ m. Similar results were observed in other injected ears that were immunolabelled ($n = 3$).



Extended Data Figure 9 | *In vivo* editing of the *Tmc1* locus from *Tmc1*^{Bth/+} ears injected with Cas9–*Tmc1*–mut3 sgRNA. A representation of the organ of Corti removed at P5 for high-throughput DNA sequencing. **a**, A confocal z-stack image showing the surface view of a dissected and labelled organ of Corti used for HTS. **b**, A cross-sectional view of the organ of Corti (along the white line in **a**) showing the positions of hair cells (MYO7A), supporting cells (SOX2) and the cells from other cochlear regions that were used for quantification. LER, lesser epithelial ridge; GER, greater epithelial ridge; SE, sensory epithelium; Lib, limbus region. DAPI-labelled nuclei are shown in blue. Quantification showed that

hair cells represented $1.45 \pm 0.05\%$ (mean \pm s.e.m., $n = 4$) of all cells in the dissected cochlea. Scale bars, 10 μm . **c**, On-target and off-target *in vivo* editing of the *Tmc1* locus in organ of Corti samples. No indels were observed at frequencies substantially above that of an untreated control sample at any of the ten off-target sites identified by GUIDE-seq (*Off-T1* to *Off-T10*). Indels were detected by HTS at the *Tmc1* on-target site and each off-target site from *in vivo* tissue samples dissected from the inner ear of neonatal mice 4 days after Cas9–*Tmc1*–mut3 RNP injection (blue), or from untreated control samples (red).

Extended Data Table 1 | Off-target editing after nucleofection of DNA plasmids encoding Cas9 and Tmc1-mut3 sgRNA into primary fibroblasts derived from *Tmc1^{Bth/Bth}* mice**a**

	5'-Sequence-3'	Mismatches (MMs)	NCBI accession	Predicted function	Location	Indels in <i>Bth/Bth</i>
<i>Bth</i>	GGGTGGGACAGAACTTCCCCAGG	0MMs	N/A		chr9	31%
<i>Off-T1</i>	GGGAGGGACAGAGCTTCCCCAGG	2MMs [4:13]	N/A		chr1	8.1%
<i>Off-T2</i>	GTGAGGGAGAGAACTTCCCCTGG	3MMs [2:4:9]	N/A		chr16	4.4%
<i>Off-T3</i>	AGTTGGTACAGAACTTCCCCAGG	3MMs [1:3:7]	NC_000068.7	CD82 antigen	chr2	2.6%
<i>Off-T4</i>	TTGTGGGACAGAAATTTCCCCAGG	3MMs [1:2:14]	N/A		chr12	3.9%
<i>Off-T5</i>	AGAGGAGACAGAACTCCCCAGG	5MMs [1:3:4:6:16]	N/A		chr13	3.4%
<i>Off-T6</i>	GGGTGGGACAGATCTTCCCCAGG	2MMs [13:20]	NC_000067.6	hemicentin-1 isoform	chr1	0.68%
<i>Off-T7</i>	GTGTAGGACAGAACTTCGCCAGG	3MMs [2:5:18]	XM_006507026.3	inositol 1,4,5-triphosphate receptor 2	chr6	1.5%
<i>Off-T8</i>	GGTGAGACCAGAGCTTCCCCTGG	6MMs [3:4:5:7:8:13]	XR_389309.3	unknown	chr5	1.2%
<i>Off-T9</i>	AGGTGGGAAGAACTTCTCCGGG	3MMs [1:9:18]	NC_000070.6	paralemmin A kinase anchor protein	chr4	1.4%
<i>Off-T10</i>	GGGTGGTAAGAACTTCTCCTGG	3MMs [7:9:18]	N/A		chr10	0.048%

b

	5'-Sequence-3'	Mismatches	NCBI accession	Location	Indels in <i>Bth/Bth</i>
<i>Bth</i>	GGGTGGGACAGAACTTCCCCAGG	0MMs		chr9	31%
<i>Off-T1'</i>	GGGAGGGACAGAGCTTCCCCAGG	2MMs [4:13]		chr1	8.1%
<i>Off-T2'</i>	GTGAGGGAGAGAACTTCCCCTGG	3MMs [2:4:9]		chr16	4.4%
<i>Off-T3'</i>	AGGAAGGCCAGAACTTCCCCTAG	4MMs [1:4:5:8]	NM_001312644.1	chr12	0.037%
<i>Off-T4'</i>	GGAGGGGGCTGAACTTCCCCAGG	4MMs [3:4:8:10]		chr9	0.071%
<i>Off-T5'</i>	GTGTGGAACAGAACTTCCCAGGG	3MMs [2:7:20]		chr1	0.046%
<i>Off-T6'</i>	CCCTGGAAACAGAACTTCCCCAAG	4MMs [1:2:3:7]		chr2	0.097%
<i>Off-T7'</i>	GCGCGGGACAGAACATCCCCTAG	3MMs [2:4:15]		chr5	0.033%
<i>Off-T8'</i>	AGGCAGGACAAACTTCCCCAAG	4MMs [1:4:5:11]		chr1	0.091%

a, Off-target sites identified by GUIDE-seq²³. Mismatch positions are indicated, counting the PAM as positions 21–23. *Off-T3*, *Off-T6*, *Off-T7*, *Off-T8* and *Off-T9* are located within predicted gene regions, while the rest are intergenic. One thousand nanograms Cas9 plasmid and 300 ng Tmc1-mut3 sgRNA plasmid were nucleofected into *Tmc1^{Bth/Bth}* fibroblasts using a LONZA 4D-Nucleofector and indels were detected by HTS at Tm *Tmc1^{Bth}* on-target and each off-target site. Mismatches compared to the on-target sequence are shown in red and PAMs are in blue. **b**, Off-target sites identified by computational prediction using the CRISPR Design Tool²⁴. Among the top eight computationally predicted off-target sites, only two (*off-T1'* and *off-T2'* with two and three mismatches, respectively) were identified as bona fide off-targets in cells by GUIDE-seq.

A metabolic function of *FGFR3–TACC3* gene fusions in cancer

Véronique Frattini^{1*}, Stefano M. Pagnotta^{1,2*}, Tala¹, Jerry J. Fan^{3,4}, Marco V. Russo¹, Sang Bae Lee¹, Luciano Garofano^{1,2,5}, Jing Zhang¹, Peiguo Shi¹, Genevieve Lewis¹, Heloise Sanson¹, Vanessa Frederick¹, Angelica M. Castano¹, Luigi Cerulo^{2,5}, Delphine C. M. Rolland⁶, Raghvendra Mall⁷, Karima Mokhtari^{8,9,10}, Kojo S. J. Elenitoba-Johnson⁶, Marc Sanson^{8,10,11}, Xi Huang^{3,4}, Michele Ceccarelli^{2,5}, Anna Lasorella^{1,12,13§} & Antonio Iavarone^{1,12,14§}

Chromosomal translocations that generate in-frame oncogenic gene fusions are notable examples of the success of targeted cancer therapies^{1–3}. We have previously described gene fusions of *FGFR3–TACC3* (F3–T3) in 3% of human glioblastoma cases⁴. Subsequent studies have reported similar frequencies of F3–T3 in many other cancers, indicating that F3–T3 is a commonly occurring fusion across all tumour types^{5,6}. F3–T3 fusions are potent oncogenes that confer sensitivity to FGFR inhibitors, but the downstream oncogenic signalling pathways remain unknown^{2,4–6}. Here we show that human tumours with F3–T3 fusions cluster within transcriptional subgroups that are characterized by the activation of mitochondrial functions. F3–T3 activates oxidative phosphorylation and mitochondrial biogenesis and induces sensitivity to inhibitors of oxidative metabolism. Phosphorylation of the phosphopeptide PIN4 is an intermediate step in the signalling pathway of the activation of mitochondrial metabolism. The F3–T3–PIN4 axis triggers the biogenesis of peroxisomes and the synthesis of new proteins. The anabolic response converges on the PGC1 α coactivator through the production of intracellular reactive oxygen species, which enables mitochondrial respiration and tumour growth. These data illustrate the oncogenic circuit engaged by F3–T3 and show that F3–T3-positive tumours rely on mitochondrial respiration, highlighting this pathway as a therapeutic opportunity for the treatment of tumours with F3–T3 fusions. We also provide insights into the genetic alterations that initiate the chain of metabolic responses that drive mitochondrial metabolism in cancer.

To investigate the transcriptional changes elicited by F3–T3, we expressed F3–T3 in immortalized human astrocytes and compared gene expression profiles of cells treated with a specific inhibitor against FGFR tyrosine kinase (TK, PD173074) or vehicle. Human astrocytes expressing F3–T3 were also compared to human astrocytes that expressed kinase-dead F3–T3 (F3–T3(K508M)) or were transduced with an empty vector (Extended Data Fig. 1a). Hierarchical clustering based on genes that were differentially expressed between F3–T3 human astrocytes and PD173074-treated F3–T3 cells showed that F3–T3 human astrocytes differed from the other three groups (Fig. 1a and Extended Data Fig. 1b). Analysis of a Gene Ontology enrichment map showed that, in addition to the expected enrichment for mitotic activity⁴, oxidative phosphorylation and mitochondrial biogenesis were the most significant categories to be enriched in F3–T3 human astrocytes for each of the three independent comparisons (Fig. 1b, Extended

Data Fig. 1c and Supplementary Table 1). We confirmed the expression changes of mitochondrial genes by quantitative PCR with reverse transcription (RT–qPCR) (Extended Data Fig. 1d).

Compared to human astrocytes expressing the empty vector or F3–T3(K508M), F3–T3 human astrocytes exhibited increased levels of mitochondrial DNA, mitochondrial mass (MitoTracker Red) and produced higher levels of ATP (Fig. 1c, d and Extended Data Fig. 1e). F3–T3 increased respiratory complex proteins (SDHB, UQCRC1 and ATP5A1) and the mitochondrial membrane transporter VDAC1 (Extended Data Fig. 1f). We also found higher levels of VDAC1 and NDUFS4 in tumours generated from mouse glioma stem cells (mGSCs) expressing human F3–T3 and small hairpin RNA (shRNA) against *Trp53* (shTrp53) (hereafter F3–T3;shTrp53) than in tumours formed by mGSCs expressing oncogenic HRAS(12V) and shTrp53 (hereafter HRAS(12V);shTrp53)^{4,7} (Extended Data Fig. 1g). Introduction of F3–T3 in human astrocytes, RPE and U251 cells increased the basal and maximal oxygen consumption rate (OCR) of these cells compared to cells transduced with F3–T3(K508M) or empty vector and this effect was reversed by FGFR-TK inhibition with AZD4547 in cells expressing exogenous F3–T3 and human glioblastoma (GBM)-derived GSC1123 cells with endogenous F3–T3⁴ (Fig. 1e and Extended Data Fig. 2a–d). F3–T3 elicited only a mild increase in the extracellular acidification rate (ECAR), leading to an increase in the OCR:ECAR ratio (Extended Data Fig. 2e, f). After treatment with the inhibitor of ATP synthase oligomycin, F3–T3 human astrocytes displayed reduced ATP levels and cell growth (by more than 70%) but were resistant to the substitution of glucose with galactose in the culture medium, a condition that imposes oxidative metabolism and markedly affected cell growth of human astrocytes treated with vector (Extended Data Fig. 2g, h). A 72-h treatment with the mitochondrial inhibitors metformin, menadione or tigecycline impaired growth of GSC1123 cells but was ineffective in GSC308 F3–T3-negative gliomaspheres⁴ (Fig. 1f). Similarly, mitochondrial inhibitors reduced the viability of F3–T3;shTrp53 mGSCs but did not affect HRAS(12V);shTrp53 mGSCs (Fig. 1g and Extended Data Fig. 2i–k). However, tigecycline decreased COX1 and COX2, two respiratory complex subunits translated by mitochondrial ribosomes⁸, and mitochondrial inhibitors reduced ATP production, indicating that these compounds were similarly active in both cell types (Extended Data Fig. 2l, m). We also found that treatment with tigecycline (50 mg kg^{–1}) suppressed tumour growth of F3–T3;shTrp53 mGSCs glioma xenografts with a more than 50% reduction in tumour

¹Institute for Cancer Genetics, Columbia University Medical Center, New York, New York 10032, USA. ²Department of Science and Technology, Università degli Studi del Sannio, Benevento 82100, Italy. ³The Arthur and Sonia Labatt Brain Tumour Research Centre, Program in Developmental and Stem Cell Biology, The Hospital for Sick Children, Toronto, Ontario M5G 1A4, Canada.

⁴Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada. ⁵BIOGEM Istituto di Ricerche Genetiche ‘G. Salvatore’, Campo Reale, 83031 Ariano Irpino, Italy.

⁶Department of Pathology and Laboratory Medicine, Perelman School of Medicine at University of Pennsylvania, Philadelphia, Pennsylvania 19104-6100, USA. ⁷Qatar Computing Research Institute (QCRI), Hamad Bin Khalifa University, Doha, Qatar. ⁸Sorbonne Universités UPMC Univ Paris 06, UMR S 1127, Inserm U 1127, CNRS UMR 7225, ICM, Paris 75013, France. ⁹AP-HP, Groupe Hospitalier Pitié Salpêtrière, Laboratoire de Neuropathologie R Escourolle, Paris 75013, France. ¹⁰Onconeurotek, AP-HP, Paris 75013, France. ¹¹AP-HP, Hôpital de la Pitié-Salpêtrière, Service de Neurologie 2, Paris 75013, France. ¹²Department of Pathology and Cell Biology, Columbia University Medical Center, New York, New York 10032, USA. ¹³Department of Pediatrics, Columbia University Medical Center, New York, New York 10032, USA. ¹⁴Department of Neurology, Columbia University Medical Center, New York, New York 10032, USA.

*These authors contributed equally to this work.

§These authors jointly supervised this work.

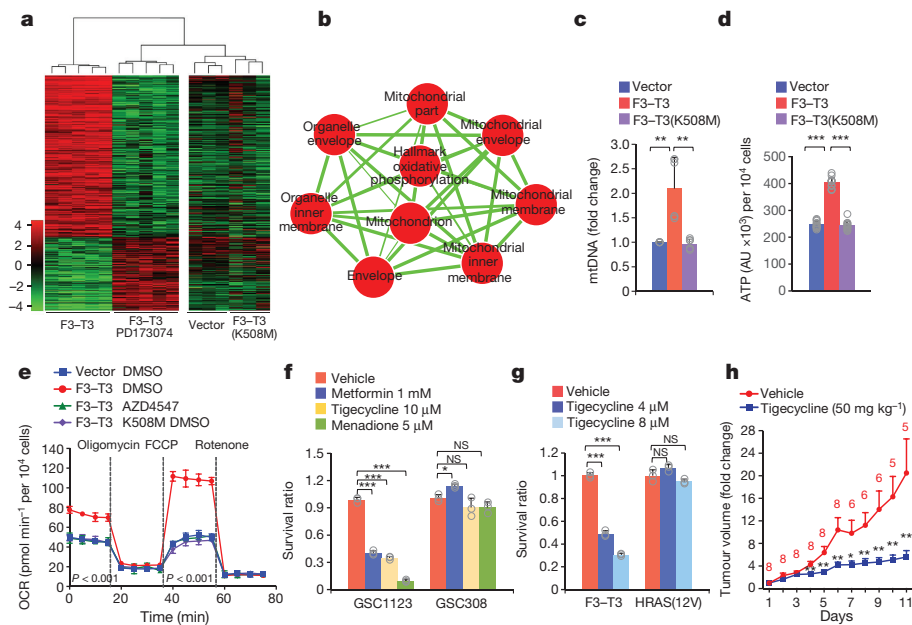


Figure 1 | Activation of mitochondrial biogenesis and metabolism by F3-T3. **a**, Hierarchical clustering of differentially expressed genes (DEGs) between F3-T3 human astrocytes, F3-T3 human astrocytes treated with PD173074 ($n = 5$ biologically independent samples per group). Human astrocytes expressing the empty vector and F3-T3(K508M) ($n = 3$ biologically independent samples per group) are included as controls. t -test $P < 0.01$ and MWW test $P < 0.01$. **b**, Enrichment map network of statistically significant GO categories ($Q < 10^{-6}$ in F3-T3 human astrocytes versus F3-T3 human astrocytes treated with PD173074 and human astrocytes expressing F3-T3(K508M) or vector). Nodes represent GO terms and lines their connectivity. Node size is proportional to the significance of enrichment and line thickness indicates the fraction of genes shared between groups. **c**, qPCR of mitochondrial DNA (mtDNA)

in human astrocytes expressing F3-T3, F3-T3(K508M) or vector. **d**, Quantification of cellular ATP in human astrocytes as in **c**. **e**, OCR of F3-T3 human astrocytes treated with or without AZD4547. **f**, Survival ratio of GSC1123 and GSC308 cells following treatment with the indicated mitochondrial inhibitors. **g**, Survival ratio of F3-T3;shTrp53 and HRAS(12V);shTrp53 mGSCs treated with vehicle or tigecycline. **h**, Tumour volume in mice treated with vehicle ($n = 8$) or tigecycline ($n = 10$). Data are fold change \pm s.e.m. of controls. The number of mice remaining in the study at each time point is indicated. Data are representative of two (**f**, **g**) or three (**e**) independent experiments. Data are fold change \pm s.d. (**c**) or mean \pm s.d. (**d**–**g**) of $n = 3$ technical replicates (**e**–**g**) or $n = 6$ (**c**) and $n = 12$ replicates (**d**) from two independent experiments. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, two-tailed t -test with unequal variance.

volume after six days, the last day on which all controls were alive. At the end of the experiment (day 11), three of the eight mice in the control group had been euthanized, whereas all mice receiving tigecycline were alive ($n = 10$; Fig. 1h and Extended Data Fig. 2n).

To identify F3-T3 substrates that drive oxidative metabolism, we performed anti-phosphorylated tyrosine (phospho-tyrosine) immunoprecipitation of tryptic digests of total cellular proteins from human astrocytes expressing F3-T3, F3-T3(K508M) or the empty vector, followed by identification of phosphopeptides by liquid chromatography–tandem mass spectrometry (Supplementary Table 2). As expected, F3-T3 showed the largest changes in phospho-tyrosine; Y647 in FGFR3 and Y684 in TACC3 showed the highest and second highest enrichment in phosphorylation, respectively. After the enrichment seen in F3-T3, the next most enriched phospho-tyrosine was Y122 of PIN4 (hereafter PIN4(Y122)), a poorly studied homologue of the cancer-driver PIN1 peptidyl-prolyl-*trans*-isomerase^{9–11} (Supplementary Table 2). This residue (Y122) is conserved in PIN4 across evolution and we found that F3-T3 interacts with endogenous PIN4 (Extended Data Fig. 3a, b). Analysis of anti-phosphotyrosine immunoprecipitations revealed that only cells expressing active F3-T3 contained tyrosine-phosphorylated PIN4, PKM2, DLG3, C1orf50 and GOLGIN84, whereas tyrosine-phosphorylated HGS was also present in FGFR3-expressing cells (Fig. 2a and Extended Data Fig. 3c, d). Treatment of GSC1123 cells with AZD4547 removed constitutive tyrosine phosphorylation of F3-T3, PIN4, PKM2, GOLGIN84 and C1orf50, whereas phospho-ERK, phospho-Stat3 and phospho-AKT were not changed (Extended Data Fig. 3e, f). We confirmed F3-T3-specific tyrosine phosphorylation of exogenous wild-type PIN4, PKM2, GOLGIN84, DLG3 and C1orf50, but phosphorylation of the corresponding un-phosphorylatable tyrosine to alanine or phenylalanine phospho-mutants was greatly reduced (Extended Data Fig. 3g). We generated and validated a

phosphorylation-specific antibody against phosphorylated PIN4(Y122) (phospho-PIN4(Y122)). The antibody detected PIN4 in cells expressing F3-T3, but not in cells transduced with vector, FGFR3 or F3-T3(K508M) (Extended Data Fig. 3h, i). Phospho-PIN4(Y122) was readily detected in F3-T3;shTrp53 mGSCs and xenografts but was absent in HRAS(12V);shTrp53 mGSCs and corresponding tumours (Extended Data Fig. 4a, b). Immunostaining of phospho-PIN4(Y122) in primary human GBM revealed that tumours with F3-T3 ($n = 14$) expressed much higher levels of phospho-PIN4(Y122) than tumours lacking F3-T3 fusions ($n = 35$, 15 of which expressed EGFR–SEPT14, a different receptor tyrosine kinase gene fusion that signals through phospho-STAT3¹²; Fig. 2b and Extended Data Fig. 4c).

Next, we expressed wild-type and the corresponding phospho-tyrosine mutants of PIN4, PKM2, DLG3, C1orf50 and GOLGIN84 in F3-T3 human astrocytes and measured oxidative metabolism. Expression of wild-type and tyrosine to alanine or phenylalanine mutants of PKM2, DLG3, C1orf50 and GOLGIN84 failed to affect the increased OCR profile of F3-T3 human astrocytes (Extended Data Fig. 4d–g). Conversely, PIN4(Y122F) but not wild-type PIN4 (PIN4(WT)) reverted basal and maximum OCR levels of F3-T3 human astrocytes to the levels of vector-expressing human astrocytes (Fig. 2c and Extended Data Fig. 4h). We observed similar effects in F3-T3 human astrocytes in which endogenous PIN4 had been silenced and replaced by the un-phosphorylatable PIN4(Y122F) phospho-mutant (Fig. 2d and Extended Data Fig. 4i). Expression of PIN4(Y122F) and PIN4(Y122A) phospho-mutants reversed the F3-T3-mediated increase in ATP (Extended Data Fig. 4j). Expression of PIN4(Y122F) also impaired soft agar clonogenicity (Fig. 2e).

To identify the gene expression signature associated with F3-T3 in human tumours, we benchmarked different statistical methods for the analysis of imbalanced datasets using synthetic data and the

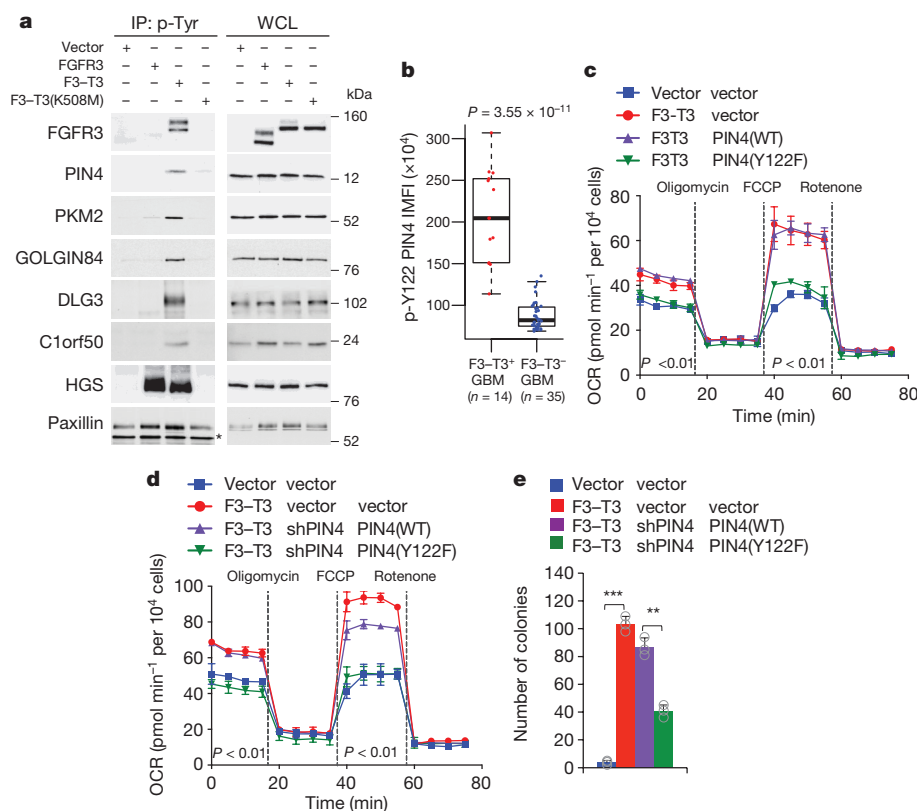


Figure 2 | Phosphorylation of PIN4 at Y122 affects mitochondrial metabolism. **a**, Immunoblot of phosphotyrosine immunoprecipitates from SF126 glioma cells (left) or whole-cell lysate (WCL, right). Paxillin is a loading control. IP, immunoprecipitation. **b**, Quantification of phospho-PIN4(Y122) integrated mean fluorescence intensity (IMFI) from F3-T3-positive and F3-T3-negative GBM. Box plot spans the first to third quartiles and whiskers show the $1.5 \times$ interquartile range. $P \leq 0.0001$, two-sided MWW test. **c**, OCR of F3-T3 human astrocytes

transduced with PIN4(WT), PIN4(Y122F) or vector. **d**, OCR of F3-T3 human astrocytes following silencing of *PIN4* and reconstitution with PIN4(WT) or PIN4(Y122F). **e**, Soft agar colony-forming assay of human astrocytes treated as in **d**. Data in **c–e** are mean \pm s.d. of one representative experiment with $n = 3$ technical replicates. Experiments were repeated three times with similar results. $**P \leq 0.01$, $***P \leq 0.001$, two-tailed *t*-test with unequal variance.

GBM transcriptome from The Cancer Genome Atlas (TCGA)¹³. The combination of the easy ensemble (ee) undersampling technique and Mann–Whitney–Wilcoxon (MWW) test statistics (ee-MWW) exhibited the best performance for correct identification of imbalanced samples and reproducible clustering (Supplementary Information and Supplementary Table 3). We used ee-MWW to generate a ranked list of genes discriminating F3-T3-positive samples in the GBM dataset of the TCGA and built a hierarchical cluster (confirmed by consensus clustering), including a small cluster of nine F3-T3-positive samples and nine fusion-like GBM (Fig. 3a, Extended Data Fig. 5a and Supplementary Table 4). The most significant biological processes enriched in F3-T3-positive GBM were mitochondrial categories (Fig. 3b and Extended Data Fig. 5b). Mitochondrial functions were also increased in fusion-like GBM, which were enriched for amplification and high expression of mitochondrial RNA polymerase (*POLRMT*, Extended Data Fig. 5c–e). Immunostaining of oxidative phosphorylation biomarkers in an independent GBM cohort revealed that F3-T3-positive tumours expressed higher levels of mitochondrial proteins (Fig. 3c and Extended Data Fig. 5f). The ee-MWW method clustered tumours with other rare oncogenes (oncogenic RAS in GBM and invasive breast carcinoma, *EGFR-SEPT14* gene fusion in GBM¹²) and identified their associated biological functions (Extended Data Fig. 5g–i and Supplementary Table 5). Using ee-MWW, we detected small and homogeneous clusters of F3-T3-positive tumours enriched for mitochondrial categories in each tumour containing recurrent F3-T3 fusions in the TCGA dataset (pan-glioma, lung squamous cell carcinoma, head and neck squamous cell carcinoma, oesophageal carcinoma, urothelial bladder carcinoma and cervical squamous cell carcinoma and endocervical adenocarcinoma; Extended Data Fig. 6a–k

and Supplementary Table 6). The transcriptional similarity of F3-T3-positive glioma was confirmed by Topological Data Analysis^{14,15} (Extended Data Fig. 6l). Finally, expression of the *F3-T3* fusion gene correlated with mitochondrial activities in the analysis of multiple cancer types (Extended Data Fig. 6m).

To identify the transcription factors that are causally related to the gene expression signature that is activated in F3-T3-positive glioma (master regulators)¹⁶, we assembled transcriptional networks from the GBM and pan-glioma datasets using the regularized gradient boosting machine algorithm that was developed for the inference of gene regulatory networks¹⁷. In both datasets, the two most active master regulators of F3-T3-positive tumours were *PPARGC1A* and *ESRRG* (encoding the PGC1 α transcriptional coactivator and the nuclear receptor ERR γ , respectively; Fig. 3d, Extended Data Fig. 6n and Supplementary Table 7). Expression of *PPARGC1A* and *ESRRG* mRNA was higher in F3-T3-positive than F3-T3-negative GBM (Extended Data Fig. 6o). Because PGC1 α is a coactivator of the oestrogen-related receptor (ERR) subfamily of nuclear receptors and acts as a master regulator of mitochondrial biogenesis and metabolism^{18,19}, we investigated whether PGC1 α and ERR γ enable the mitochondrial functions induced by F3-T3. Introduction of F3-T3 in human astrocytes expressing PIN4(WT) increased *PPARGC1A* mRNA and PGC1 α protein and the expression of genes involved in reactive oxygen species (ROS) detoxification²⁰ (Extended Data Fig. 7a–d and Supplementary Table 7). Accordingly, PGC1 α accumulated at higher levels in F3-T3-positive GSC1123 cells and F3-T3;shTrp53 mGSCs compared to F3-T3-negative GSC308 cells and HRAS(12V);shTrp53 mGSCs, respectively (Fig. 3e). However, replacement of PIN4 with the un-phosphorylatable Y122F mutant PIN4(Y122F) blunted

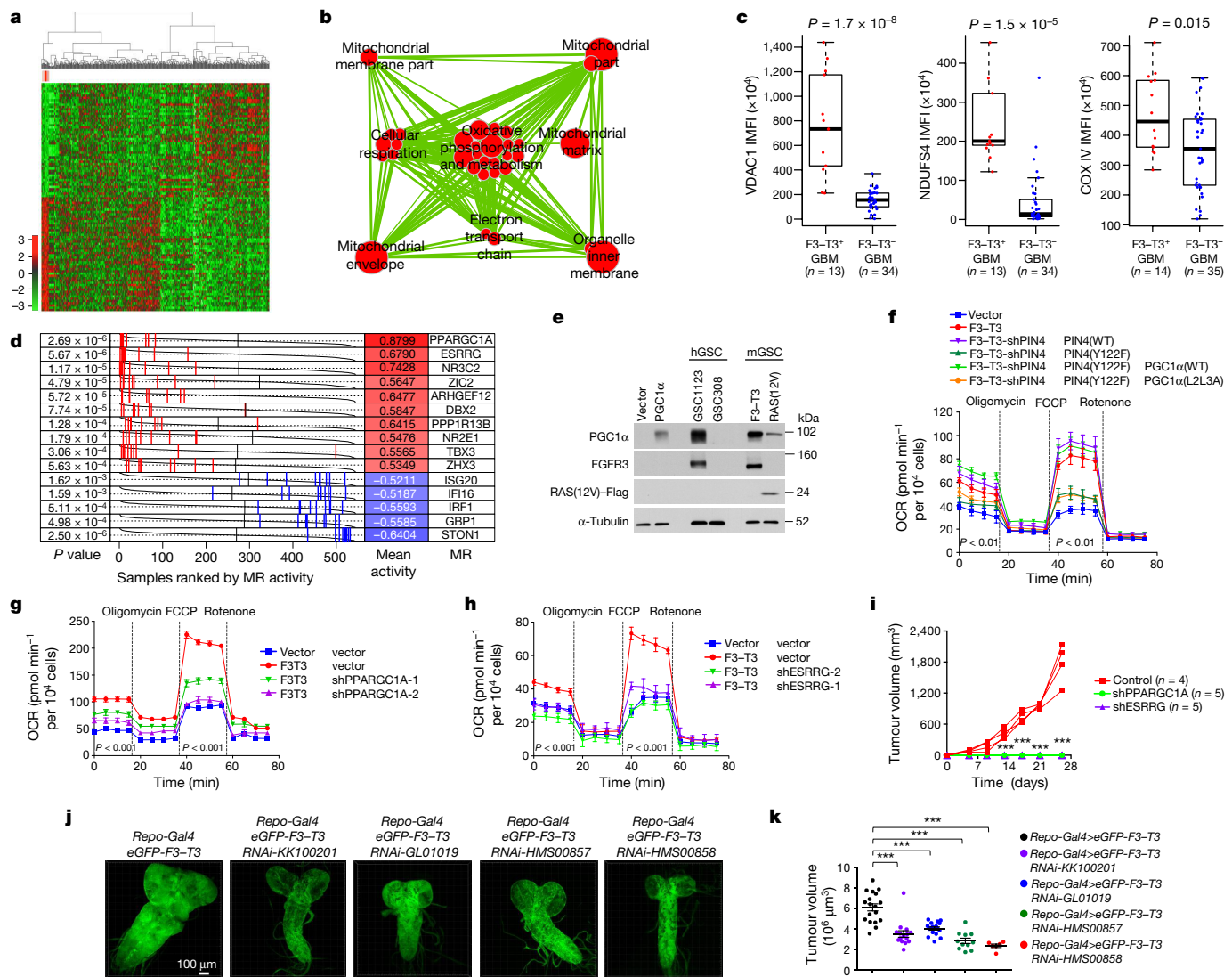


Figure 3 | PGC1 α and ERR γ are required for F3-T3-mediated mitochondrial metabolism and tumorigenesis. **a**, Hierarchical clustering of GBM ($n = 534$) and normal brain ($n = 10$) from the TCGA using DEGs in nine F3-T3-positive samples (red) versus the F3-T3-negative samples. **b**, Enrichment map network of statistically significant GO categories in nine F3-T3-positive samples (upper-tailed MWW-GST $Q < 0.001$, normalized enrichment score (NES) > 0.6). Nodes represent GO terms and lines their connectivity. Node size is proportional to number of genes in the GO category and line thickness indicates the fraction of genes shared between groups. **c**, Quantification of IMFI of VDAC1, NDUFS4 and COXIV in F3-T3-positive and F3-T3-negative GBM. Box plot spans the first to third quartiles and whiskers show the $1.5 \times$ interquartile range. $P \leq 0.0001$ (VDAC and NDUFS4); $P \leq 0.05$ (COXIV), two-sided MWW test. **d**, Master regulator (MR) activity in GBM. Grey curves represent the activity of each master regulator. Red or blue lines indicate individual F3-T3-positive GBM displaying high or low master regulator activity, respectively ($n = 534$). P value, two-sided MWW test for differential activity (left) and mean of the activity (right) of the master regulator in F3-T3-positive versus F3-T3-negative samples are indicated.

F3-T3-mediated induction of PGC1 α (Extended Data Fig. 7a, c). The inhibition of mitochondrial metabolism and reduction in soft agar clonogenicity by PIN4(Y122F) in F3-T3 human astrocytes were both rescued by overexpression of PGC1 α (WT). Conversely, PGC1 α (L2L3A), which contains mutations in the nuclear receptor boxes L2 and L3 that are critical for binding ERR γ ¹⁸ could not rescue F3-T3-mediated activation of mitochondrial metabolism in F3-T3 human astrocytes expressing PIN4(Y122F) (Fig. 3f and Extended

Data Fig. 7e, f). Finally, loss of PGC1 α by shRNA and CRISPR-Cas9 gene editing reversed the activation of mitochondrial respiration by F3-T3 and depletion of ERR γ produced similar effects (Fig. 3g, h and Extended Data Fig. 7g–m). PGC1 α silencing inhibited soft agar colony formation by F3-T3 human astrocytes and impaired self-renewal of GSC1123 cells (Extended Data Fig. 7n–p). Silencing of either *PPARGC1A* or *ESRRG* prevented tumour xenograft formation of F3-T3 human astrocytes in mice (Fig. 3i and Extended Data

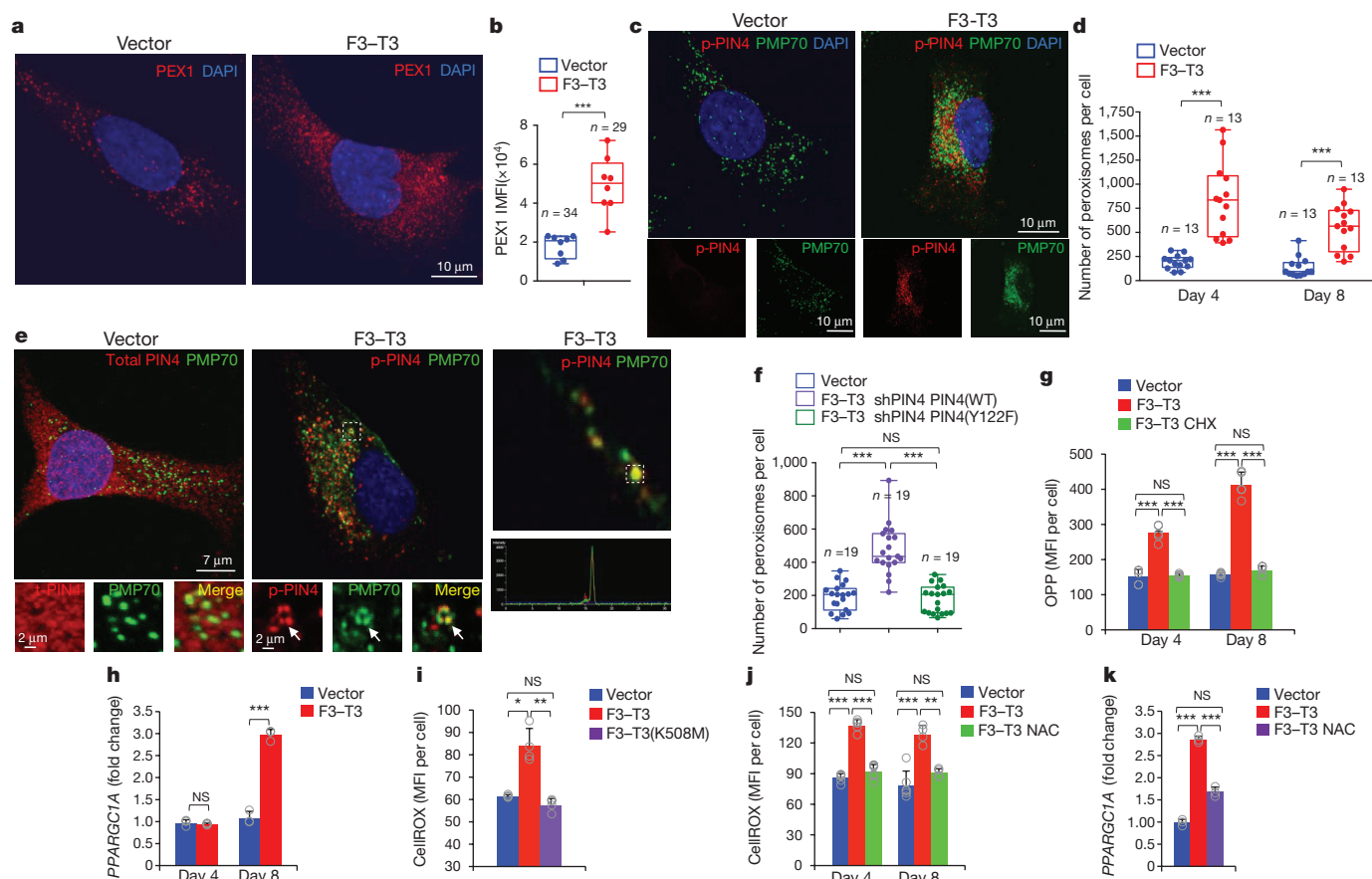


Figure 4 | Expression of F3-T3 fusion induces peroxisome biogenesis through phosphorylation of PIN4(Y122). **a**, Representative confocal micrographs of PEX1 immunostaining (red) and DAPI (blue) in vector and F3-T3-expressing human astrocytes. **b**, Quantification of PEX1 IMFI in samples stained as in **a** ($n = 34$ and 26 cells for vector and F3-T3, respectively). **c**, Representative confocal micrographs of double (top) and single (bottom) immunostaining for phospho-PIN4(Y122) (p-PIN4, red) and PMP70 (green) in vector and F3-T3-expressing human astrocytes. **d**, Quantification of peroxisome number per cell 4 and 8 days after F3-T3 expression in human astrocytes ($n = 13$ cells). **e**, Representative confocal micrographs of double immunostaining for total PIN4 (t-PIN4) and PMP70 in F3-T3-expressing (middle) human astrocytes. Arrowheads, phospho-PIN4(Y122)-PMP70 co-localization (bottom). Right, phospho-PIN4(Y122)-PMP70 co-localization (top) with corresponding spectral intensity profile (bottom); co-localization coefficients: Pearson's correlation $r = 0.935963$; Mander's overlap = 0.959905 ; Mander's overlap coefficients $k_1 = 0.934640$, $k_2 = 0.985853$; colocalization coefficients $c_1 = 1.000000$, $c_2 = 0.999792$. **f**, Quantification of peroxisome number per

cell in F3-T3 human astrocytes following silencing of *PIN4* and expression of *PIN4*(WT) or *PIN4*(Y122F) ($n = 19$ cells). **g**, Quantification of protein biosynthesis by O-propargyl-puromycin (OPP) incorporation in human astrocytes that were treated as in **d**. Cycloheximide (CHX)-treated cultures are used as negative controls ($n = 5$ technical replicates). **h**, RT-qPCR of *PPARGC1A* in human astrocytes that were treated as in **d** ($n = 3$ technical replicates). **i**, Quantification of cellular ROS in vector-, F3-T3- and F3-T3-expressing human astrocytes. Bar graphs from one representative experiment ($n = 4$ technical replicates). **j**, Analysis of cellular ROS in human astrocytes that were treated as in **d**. Bar graphs from one representative experiment ($n = 5$ or 6 technical replicates). **k**, RT-qPCR of *PPARGC1A* in vector- or F3-T3-expressing cells that were treated with vehicle or *N*-acetyl-L-cysteine (NAC) ($n = 3$ technical replicates). Data are mean \pm s.d. (**g-k**). *P*-values: **b**, **d**, **f**, two-sided MWW test; **g-k**, two-tailed *t*-test with unequal variance. Box plots span the first to third quartiles and whiskers indicate the smallest and largest values. Experiment in **i** was repeated twice; all other experiments were repeated three times with similar results.

Fig. 7q) and impaired *in vivo* tumour growth of F3-T3;shTrp53 mGSCs but not HRAS(12V)shTrp53 mGSCs (Extended Data Fig. 7r-v). Next, we developed a brain tumour model in *Drosophila* by ectopically expressing human F3-T3 using the glial-specific driver *repo-Gal4* (ref. 21). *repo-Gal4*-F3-T3 transgenic flies manifested glial neoplasia with enlargement and malformation of the larval brain lobe and ventral nerve cord, leading to larval lethality (Extended Data Fig. 8a, b). Cell number and proliferation of glial cells were enhanced in *repo-Gal4*-F3-T3 flies (Extended Data Fig. 8c, d). Cell-autonomous RNA interference (RNAi)-mediated knockdown of *spargel* (*srl*, the *Drosophila* orthologue of *PPARGC1A*)²² in *repo-Gal4*-F3-T3 flies reduced glial tumour volume and decreased the number of Repo⁺ glial cells, without affecting *repo-Gal4*-driven F3-T3 protein expression in F3-T3-expressing flies or normal brain development in wild-type animals without F3-T3 (Fig. 3j, k and Extended Data Figs 8e-g, 9a-d). *srl* knockdown did not rescue *Repo-Gal4*-F3-T3 animals to adult

viability, confirming that suppressors of glial neoplasia in *Drosophila* are infrequent rescuers of organismic lethality²³ (Extended Data Fig. 9e). *srl* knockdown in a *Drosophila* model of glioma driven by constitutively active EGFR (dEGFR^Δ) and PI3K (Dp110^{CAAX}) in the glial lineage²⁴ resulted in minor to no effects on tumour volume, thus highlighting the specific sensitivity of F3-T3 tumorigenesis to the perturbation of *srl* expression (Extended Data Fig. 9f, g). RNAi-mediated knockdown of the *Drosophila* oestrogen-related receptor (*ERR*) also reduced F3-T3 glial tumour volume (Extended Data Fig. 9h, i).

To determine the mechanism by which phospho-PIN4(Y122) mediates F3-T3 signalling, we studied the subcellular compartmentalization of PIN4 and phospho-PIN4(Y122) and sought to uncover the set of cellular proteins interacting with PIN4. Unphosphorylated PIN4 was diffusely localized in the cytoplasm and nuclear membrane whereas phospho-PIN4(Y122) was concentrated in larger cytoplasmic vesicle-like structures that co-localized with F3-T3 (Extended Data

Figs 3h, 10a, b). Using mass spectrometry analysis of PIN4 immunofluorescence complexes, we found that the peroxisomal biogenesis complex²⁵ formed by PEX1 and PEX6 is the top ranking PIN4 interactor. Other PIN4-associated proteins are implicated in vesicle formation and trafficking, nuclear and mitochondrial RNA metabolism and translation, ribosomal activity and nuclear pore/envelope functions (Extended Data Fig. 10c, d and Supplementary Table 8). Quantitative immunofluorescence revealed that PEX1 increased 2.7-fold in F3–T3 human astrocytes without changes in *PEX1* mRNA (Fig. 4a, b and Extended Data Fig. 10e, f). To investigate whether F3–T3 signals through phospho-PIN4(Y122) to promote peroxisome biogenesis, we acutely transduced human astrocytes with a F3–T3-expressing lentivirus, and found that both phospho-PIN4(Y122) and the total number of PMP70-positive peroxisomes were increased after four days (4.3 fold increase in peroxisomes per cell; Fig. 4c, d and Extended Data Fig. 10g). Phospho-PIN4(Y122)-positive cytoplasmic structures in F3–T3 human astrocytes, but not unphosphorylated PIN4 in vector-transduced human astrocytes, colocalized with PMP70, indicating that phospho-PIN4(Y122) trafficks to new peroxisomal membranes (Fig. 4e). Increased peroxisome biogenesis induced by acute expression of F3–T3 was prevented when F3–T3 was introduced in cells in which PIN4 had been replaced by the unphosphorylatable PIN4(Y122F) mutant (Fig. 4f). F3–T3 also induced a phospho-PIN4(Y122)-dependent early increase in new protein synthesis (Fig. 4g and Extended Data Fig. 10h). Conversely, PGC1 α and mitochondrial gene expression were unchanged four days after acute expression of F3–T3 but increased after eight days (Fig. 4h and Extended Data Fig. 10i). Peroxisome biogenesis and new protein synthesis can both generate ROS, and ROS are crucial inducers of PGC1 α ^{20,26,27}. F3–T3 but not F3–T3(K508M) increased ROS at the four-day time point and this effect required PIN4(Y122) phosphorylation (Fig. 4i, j and Extended Data Fig. 10j). Treatment of F3–T3 human astrocytes with the ROS inhibitor *N*-acetyl-L-cysteine eliminated approximately 70% of the increase in PGC1 α induced by F3–T3 (Fig. 4k), thus indicating that ROS are responsible for most of the increase in PGC1 α induced by F3–T3.

In conclusion, we describe, using an integrated computational and experimental framework, the chain of events propagated by F3–T3 in cancer. Signalling through phospho-PIN4(Y122) triggers vesicle trafficking to deliver building blocks for biogenesis of peroxisomes and new protein synthesis. The coordinated activation of these anabolic pathways results in the accumulation of ROS, which in turn increases PGC1 α -ERR γ and mitochondrial metabolism. Thus, rather than impinging exclusively on mitochondrial circuits, the oncogenic signals driving mitochondrial respiration operate within larger contexts of anabolic effectors. Dependency on mitochondrial metabolism of GBM with F3–T3 suggests that inhibitors of oxidative phosphorylation may be beneficial for patients with F3–T3-positive tumours.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 December 2016; accepted 24 November 2017.

Published online 3 January 2018.

- Ali, M. A. Chronic myeloid leukemia in the era of tyrosine kinase inhibitors: an evolving paradigm of molecularly targeted therapy. *Mol. Diagn. Ther.* **20**, 315–333 (2016).
- Di Stefano, A. L. *et al.* Detection, characterization, and inhibition of FGFR-TACC fusions in IDH wild-type glioma. *Clin. Cancer Res.* **21**, 3307–3317 (2015).
- Gerber, D. E. & Minna, J. D. ALK inhibition for non-small cell lung cancer: from discovery to therapy in record time. *Cancer Cell* **18**, 548–551 (2010).
- Singh, D. *et al.* Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science* **337**, 1231–1235 (2012).
- Costa, R. *et al.* FGFR3-TACC3 fusion in solid tumors: mini review. *Oncotarget* **7**, 55924–55938 (2016).
- Lasorella, A., Sanson, M. & Iavarone, A. FGFR-TACC gene fusions in human glioma. *Neuro-oncol.* **19**, 475–483 (2017).
- Niola, F. *et al.* Mesenchymal high-grade glioma is maintained by the ID-RAP1 axis. *J. Clin. Invest.* **123**, 405–417 (2013).

- Škrtić, M. *et al.* Inhibition of mitochondrial translation as a therapeutic strategy for human acute myeloid leukemia. *Cancer Cell* **20**, 674–688 (2011).
- Uchida, T., Fujimori, F., Tradler, T., Fischer, G. & Rahfeld, J. U. Identification and characterization of a 14 kDa human protein as a novel parvulin-like peptidyl prolyl *cis/trans* isomerase. *FEBS Lett.* **446**, 278–282 (1999).
- Uchida, T. *et al.* Pin1 and Par14 peptidyl prolyl isomerase inhibitors block cell proliferation. *Chem. Biol.* **10**, 15–24 (2003).
- Yeh, E. S. & Means, A. R. PIN1, the cell cycle and cancer. *Nat. Rev. Cancer* **7**, 381–388 (2007).
- Frattini, V. *et al.* The integrated landscape of driver genomic alterations in glioblastoma. *Nat. Genet.* **45**, 1141–1149 (2013).
- He, H. & Garcia, E. A. Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.* **21**, 1263–1284 (2009).
- Carlsson, G. Topology and data. *Bull. Am. Math. Soc.* **46**, 255–308 (2009).
- Nicolau, M., Levine, A. J. & Carlsson, G. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proc. Natl Acad. Sci. USA* **108**, 7265–7270 (2011).
- Carro, M. S. *et al.* The transcriptional network for mesenchymal transformation of brain tumours. *Nature* **463**, 318–325 (2010).
- Mall, R. *et al.* RGBM: regularized gradient boosting machines for the identification of transcriptional regulators of discrete glioma subtypes. Preprint at <https://doi.org/10.1101/132670> (2017).
- Devarakonda, S. *et al.* Disorder-to-order transition underlies the structural basis for the assembly of a transcriptionally active PGC-1 α /ERR γ complex. *Proc. Natl Acad. Sci. USA* **108**, 18678–18683 (2011).
- Lin, J., Handschin, C. & Spiegelman, B. M. Metabolic control through the PGC-1 family of transcription coactivators. *Cell Metab.* **1**, 361–370 (2005).
- St-Pierre, J. *et al.* Suppression of reactive oxygen species and neurodegeneration by the PGC-1 transcriptional coactivators. *Cell* **127**, 397–408 (2006).
- Xiong, W. C., Okano, H., Patel, N. H., Blendy, J. A. & Montell, C. repo encodes a glial-specific homeo domain protein required in the *Drosophila* nervous system. *Genes Dev.* **8**, 981–994 (1994).
- Tiefenböck, S. K., Baltzer, C., Egli, N. A. & Frei, C. The *Drosophila* PGC-1 homologue Spargel coordinates mitochondrial activity to insulin signalling. *EMBO J.* **29**, 171–183 (2010).
- Read, R. D. *et al.* A kinome-wide RNAi screen in *Drosophila* glia reveals that the RIO kinases mediate cell proliferation and survival through TORC2-Akt signaling in glioblastoma. *PLoS Genet.* **9**, e1003253 (2013).
- Read, R. D., Cavenee, W. K., Furnari, F. B. & Thomas, J. B. A *Drosophila* model for EGFR-Ras and PI3K-dependent human glioma. *PLoS Genet.* **5**, e1000374 (2009).
- Smith, J. J. & Aitchison, J. D. Peroxisomes take shape. *Nat. Rev. Mol. Cell Biol.* **14**, 803–817 (2013).
- Boveris, A., Oshino, N. & Chance, B. The cellular production of hydrogen peroxide. *Biochem. J.* **128**, 617–630 (1972).
- Han, J. *et al.* ER-stress-induced transcriptional regulation increases protein synthesis leading to cell death. *Nat. Cell Biol.* **15**, 481–490 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank C. Scuoppo for donation of the pLCiG2 plasmid and support with the gRNA design, E. Chen for identification of PIN4 immunocomplexes and H. Li for high-content microscopy. This work was supported by NIH R01CA101644, U54CA193313 and R01CA131126 to A.L.; R01CA178546, U54CA193313, R01CA179044, R01CA190891, R01NS061776 and The Chemotherapy Foundation to A.L.; SickKids Garrison Family Cancer Centre Pitblado Discovery and Ontario Institute for Cancer Research (OICR) Brain Translational Research Initiative to X.H.; American Brain Tumor Association (ABTA) and a Cancer Biology Training Grant (T32CA009503) fellowship to V.Fra.; a NRF-2013R1A6A3A03063888 fellowship to S.B.L.; an Italian Association for Cancer Research (AIRC) fellowship to M.V.R.

Author Contributions A.I. and A.L. conceived and coordinated the studies and provided overall supervision. M.C. and S.M.P. developed and performed bioinformatics analyses and wrote the computational sections. V.Fra. performed cell, molecular biology and metabolic assays, with help of T., M.V.R., A.M.C. and S.B.L. J.J.F. and X.H. developed and analysed the *Drosophila* F3–T3 model. K.S.J.E.-J. and D.M.C.R. conducted the phosphoproteomics experiments. M.S. and K.M. provided GBM tissues and assisted with immunostaining. G.L., T., V.Fra. and H.S. performed immunostaining and protein analyses. T. and P.S. performed mouse experiments. L.G., J.Z., L.C. and R.M. conducted gene expression and bioinformatics analyses. A.I. and A.L. wrote the manuscript with input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to A.L. (al2179@cumc.columbia.edu) or A.I. (ai2102@cumc.columbia.edu).

Reviewer Information Nature thanks R. Cagan, P. Mischel, M. Ochs and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Datasets. Tumour sample data are from the TCGA. Details about the cohorts and analysed samples can be found in the Supplementary Information.

Resampling methods for ranked list generation in imbalanced datasets. Preliminary testing and the ee-MWW values are reported in the Supplementary Information.

Gene Ontology networks. Gene Ontology (GO) enrichment was computed using MWW test statistics for the genes positively regulated in tumours with *FGFR3-TACC3* or other genetic alterations of interest (for example, *RAS* and *EGFR-SEPT14*). The significant GO terms from MWW-gene set test (GST) analysis (Supplementary Information) were further analysed using the Enrichment Map²⁸ application of Cytoscape²⁹. In the network, nodes represent the terms and edges represent known term interactions and are defined by the number of shared genes between the pair of terms. Size of the nodes is proportional to statistical significance of the enrichment (Fig. 1b and Extended Data Fig. 1c) or the number of genes in the category (Fig. 3b and Extended Data Figs 5c, 6c, f, h). The overlap between gene sets is computed according to the overlap coefficient (OC), defined as:

$$OC = \frac{|A \cap B|}{\min(|A|, |B|)}$$

where A and B are two gene sets, and $|X|$ equals to the number of elements within set X ³⁰. We set a cutoff of $OC > 0.5$ to select the overlapping gene sets.

Correlation analysis between GO NES and the expression of F3–T3. We selected 19 human samples with F3–T3 fusions from ref. 31 and the TCGA fusion gene Data Portal³². Starting from fastq data, we applied the ChimeraScan pipeline³³ to compute the total number of reads supporting the fusion (Supplementary Table 6m). From TCGA, we obtained the legacy level 3 RNA sequencing by expectation maximization (RSEM) counts of the samples. By using the EDASeq methodology³⁴, we corrected the counts for GC content and applied full-quantile normalization. We transformed the normalized counts in the transcripts per million abundance quantification, applied MWW-GST to each sample and collected the NES. We used the MDSigDB collections c5.bp, c5.mf, c5.cc and hallmark collections of gene sets. We compared each gene set with the number of reads supporting the F3–T3 fusions by using the Spearman's rank correlation index (Supplementary Table 6n). To test the correlation, we assumed the alternative hypothesis of the correlation greater than zero.

Assembly of the transcriptional interactomes. To identify master regulators of the gene expression signature activated in the F3–T3-positive glioma subgroup, we first assembled independent transcriptional networks from gene expression profiles of GBM and pan-glioma datasets using the regularized gradient boosting machine algorithm (RGM) (package available from CRAN at <https://cran.r-project.org/web/packages/RGM/index.html>). RGM was used to identify regulators of the molecular subtypes of brain tumours^{17,35}. We used gene expression profiles and a predefined list of 2,137 gene regulators or transcription factors (master regulators) as input. This process was independently applied to obtain GBM and pan-glioma transcriptional interactomes comprising 430,104 (median regulon size: 203) and 300,969 (median regulon size: 141) transcriptional interactions, respectively, of which 188,238 were overlapping.

Master regulator activity. To identify the master regulators of the gene expression signature activated in F3–T3-positive glioma, we modified a method that we had previously described¹⁶. In brief, the activity of a master regulator MR, defined as the index that quantifies the activation of the transcriptional program of that specific master regulator in each sample S_i , is calculated as follows:

$$Act(S_i, MR) = \frac{1}{N} \sum_{k=1}^N t_{ki}^+ - \frac{1}{M} \sum_{j=1}^M t_{ji}^-$$

where t_{ki}^+ is the expression level of the k -th positive target of the master regulator in the i -th sample, t_{ji}^- is the expression level of the j -th negative target of the master regulator in the i -th sample, N (or M) the number of positive (or negative) targets present in the regulon of the considered master regulator. If $Act(S_i, MR) > 0$, the master regulator is activated in that particular sample, if $Act(S_i, MR) < 0$, the master regulator is inversely activated, if $Act(S_i, MR) \approx 0$, it is deactivated. We used the MWW test to select master regulators that showed a significant difference between the F3–T3-positive samples and all the other samples. In Supplementary Table 7a, b, we present the list of master regulators obtained by applying master regulators analysis ($\left| \log_2 \left(\frac{NES}{1 - NES} \right) \right| > 2.0$) and significance of differential activity < 0.01 .

Topological data analysis. Topological data analysis^{14,15} (TDA) of the pan-glioma dataset was based on the Mapper algorithm³⁶. The topological network was built using the Ayasdi platform (<http://www.ayasdi.com>). Several open-source implementations of Mapper are available (<https://github.com/MLWave/kepler-mapper>, <http://danifold.net/mapper/>, <https://github.com/RabadanLab/sakmapper>,

<https://github.com/paultpearson/TDAmapper>). TDA was performed using the expression matrix of the top 100 genes differentially expressed between F3–T3-positive tumours and the remaining tumours as shown in Extended Data Fig. 6a. Mapper uses a dimensionality reduction algorithm and produces a topological representation of the data that preserves locality. The projection space of the dimensional-reduction algorithm is covered with overlapping bins. The data points that fall in each bin are then clustered in the original high-dimensional space. A network is constructed by assigning a node to each cluster, and clusters that share one or more samples are connected by an edge. The result is a low-dimensional network representation of the data in which nodes represent sets of samples with similar global transcriptional profiles, and edges connect nodes that have at least one sample in common. For our analysis we used 2D Locally Linear Embedding³⁷ as dimensional-reduction algorithm and variance normalized Euclidean metric³⁸ as distance. Single-linkage clustering was performed in each of the pre-images of the bins using a previously described algorithm³⁹. The number of bins (resolution) for each dimension was 20 and the degree of overlap (gain) between neighbouring bins was 66%. The size of the bin was chosen such that the number of samples in each row or column of bins was the same. The open-source implementations of Mapper produce results consistent with those obtained from the Ayasdi platform⁴⁰.

Transcriptomic analysis of human astrocytes. We performed comparative analysis of gene expression of human astrocytes transduced with a lentivirus expressing F3–T3 treated with vehicle (F3–T3 and DMSO, $n = 5$ replicates), F3–T3 treated with the FGFR inhibitor PD173074 for 12 h (F3–T3 and PD173074, $n = 5$ replicates), F3–T3(K508M) treated with vehicle (F3–T3(K508M) and DMSO, $n = 3$ replicates) and empty vector treated with vehicle (vector DMSO, $n = 3$ replicates). Expression data were obtained using the Illumina human HT12v4 gene expression array. The list of 4,034 differentially expressed genes between the F3–T3 and DMSO and F3–T3 and PD173074 groups (t -test $P < 0.01$ and MWW test $P < 0.01$) was used to construct a heat map comprising the whole dataset in which vector and DMSO and F3–T3(K508M) and DMSO are control groups. Samples were clustered using the hierarchical clustering algorithm based on the Ward linkage method and Euclidean distance as implemented in R. Finally, the GO enrichment analysis was performed using the ranked list obtained from three independent comparisons: F3–T3 versus F3–T3 treated with PD173074; F3–T3 versus F3–T3(K508M); F3–T3 versus vector using the Java version of GSEA. For each comparison, statistically significant GO terms with $Q < 10^{-6}$ were selected. The statistically significant pathways common to all three comparisons were included in the construction of the visual network using the Enrichment Map application²⁸ of Cytoscape²⁹. The microarray data have been deposited in ArrayExpress with accession number E-MTAB-6037.

Identification of proteins phosphorylated by the F3–T3 gene fusion using mass spectrometry. Cells were lysed in buffer containing 9 M urea, 20 mM HEPES pH 8.0, 0.1% SDS and a cocktail of phosphatase inhibitors. Six milligrams of protein were reduced with 4.5 mM DTT, alkylated with 10 mM iodoacetamide and digested with trypsin overnight at 37 °C. Samples were desalted on a C18 cartridge (Sep-Pak plus C18 cartridge, Waters). Each sample was prepared in triplicate. Phosphopeptide enrichments were performed as previously described⁴¹. An LTQ Orbitrap XL (ThermoFisher) in-line with a Paradigm MS2 HPLC (Michrom bioresources) was used to acquire high-resolution mass spectrometry and tandem mass spectrometry data. Technical duplicate data for each of the metal-oxide affinity chromatography elutions and triplicate data for the phosphotyrosine immunoprecipitation samples were acquired.

RAW files were converted to mzXML using msconvert⁴² and searched against the Swissprot Human protein database (9 January 2013 release) appended with common proteomics contaminants and reverse sequences as decoys. Searches were performed with X!Tandem (version 2010.10.01.1) using the k -score plugin^{43,44}. For all searches the following search parameters were used: parent monoisotopic mass error of 50 parts per million (p.p.m.); fragment ion error of 0.8 daltons; allowing for up to two missed tryptic cleavages. Variable modifications were oxidation of methionine (+15.9949@M), carbamidomethylation of cysteine (+57.0214@C), and phosphorylation of serine, threonine, and tyrosine (+79.9663@[STY]). The search results were then post-processed using PeptideProphet and ProteinProphet^{45–47}. Spectral counts were obtained for each cell line using ABACUS⁴⁸. Immunoprecipitation data of phospho-tyrosine enrichment were processed through ABACUS separately from the MOAC enrichment data. ABACUS results were filtered to only retain proteins with a ProteinProphet probability > 0.7 . Only phosphorylated peptides with a probability > 0.8 were considered for spectral counting. For tyrosine enrichment these ABACUS parameters resulted in a protein false discovery rate (FDR) of 0.0045. This ABACUS output was used for all subsequent analysis to quantify the relative abundance of phosphorylated peptides or proteins. Phospho-site localization was performed with an in-house reimplementation of the Ascore algorithm as previously described⁴⁹. Ascore values represent the probability of detection owing to chance, with scores

>19 corresponding to sites localized with >99% certainty. From four biological replicates, the application of stringent criteria selected 22 top-scoring candidate substrates of F3–T3 that exhibited at least a 1.5-fold enrichment in F3–T3 human astrocytes compared to human astrocytes expressing F3–T3(K508M) or the empty vector (Supplementary Table 2).

Identification of PIN4 complexes by mass spectrometry. Endogenous cellular PIN4 complexes were purified from the cell line H1299 transduced with the F3–T3-expressing lentivirus. Cellular lysates were prepared in 50 mM Tris-HCl, 250 mM NaCl, 0.2% NP40, 1 mM EDTA, 10% glycerol, protease and phosphatase inhibitors. PIN4 and mock immunoprecipitates were recovered with PIN4 antibody (Abcam, ab155283) and rabbit IgG, respectively. Immunocomplexes captured on protein A/G agarose beads were washed with lysis buffer containing 300 mM NaCl and 0.3% NP40. Bound polypeptides were eluted with the PIN4 peptide used as the epitope for the PIN4 antibody (KPVFTDPPVKTKFGYH, Abcam, ab155283). Eluates were run on SDS–PAGE gels and four gel slices were cut from the lane containing PIN4 immunoprecipitates (columns A1, B1, C1, D1; Supplementary Table 8). Four similar gel slices were cut from the lane containing control rabbit IgG immunoprecipitates (columns A2, B2, C2, D2; Supplementary Table 8). The excised gel pieces were rehydrated and digested in 80 µl of 12.5 ng µl^{−1} Trypsin Gold and 50 mM ammonium bicarbonate at 37°C overnight. Extracted peptides were dried, reconstituted in 30 µl 0.1% TFA and stored at −20°C before analysis.

The concentrated peptide mix was reconstituted in a solution of 2% acetonitrile (ACN), 2% formic acid (FA) and eluted from the column using a Dionex Ultimate 3000 Nano LC system. The application of a 2.0-kV distal voltage electrosprayed the eluting peptides directly into the Thermo Fusion Tribrid mass spectrometer equipped with an EASY-Spray source (Thermo Scientific). Mass spectrometer-scanning functions and HPLC gradients were controlled by the Xcalibur data system (Thermo Finnigan). Tandem mass spectra from raw files were searched against a human protein database using the Proteome Discoverer 1.4 software (Thermo Finnigan). The peptide mass search tolerance was set to 10 p.p.m. A minimum sequence length of seven amino acids residues was required. Only fully tryptic peptides were considered. Spectral counts were used for estimation of relative protein abundance between samples analysed directly on long gradient reverse phase liquid chromatography–tandem mass spectrometry. A specificity score of proteins interacting with PIN4 was computed for each polypeptide as described⁵⁰. In brief, we compared the number of peptides identified from our mass spectrometry analysis to those reported in the CRAPome database that includes a list of potential contaminants from affinity purification–mass spectrometry experiments (<http://www.crapome.org/>). The specificity score is computed as $(p \times c) / S_{av} \times S_{max} \times E$, where p is the identified peptide count; c the cross-correlation score for all candidate peptides queried from the database; S_{av} the averaged spectral counts from CRAPome; S_{max} the maximal spectral counts from CRAPome; and E the total number of experiments that were found in the CRAPome database.

Cell culture. *Human cell lines.* h-TERT-immortalized human astrocytes⁵¹, SF126 cells⁵², U87 (ATCC HTB-14), h-TERT-RPE-1 (ATCC CRL-4000), HEK293T (ATCC CRL-11268), U251 (Sigma 09063001) cells. Cell lines were cultured in DMEM supplemented with 10% fetal bovine serum (FBS, Sigma). Cells were transfected using Lipofectamine 2000 (Invitrogen) or the calcium phosphate method. *Mouse glioma stem cells.* F3–T3;shTrp53 and HRAS(12V);shTrp53 mGSCs were isolated from the brains of mice that had received injection of lentivirus containing a bi-cistronic expression cassette including F3–T3 or HRAS(12V) and *Trp53* shRNA into the dentate gyrus as described^{4,7}. Mice showing neurological symptoms were euthanized 2–4 months after intracranial injection, and brain tumours were identified macroscopically, dissected and cultured in DMEM:F12 containing 1 × N2 and B27 supplements (Invitrogen) and human recombinant FGF2 and EGF (20 ng ml^{−1} each; Peprotech). Studies were approved by the IACUC at Columbia University (AAAL7600).

Human glioma stem cells. The GBM-derived glioma stem cells (GSCs) used in this study have been described previously^{4,12}. GBM-derived GSCs were grown in DMEM:F12 containing 1 × N2 and B27 supplements (Invitrogen) and human recombinant FGF2 and EGF (20 ng ml^{−1} each; Peprotech). Cells were transduced using lentiviral particles in medium containing 4 µg ml^{−1} of polybrene (Sigma). Cells were routinely tested for mycoplasma contamination using the Mycoplasma Plus PCR Primer Set (Agilent Technologies) and were found to be negative. Cell authentication was performed using short-tandem repeats (STR) at the ATCC facility.

Limiting dilution assay (LDA) for human GSCs was performed as described previously⁵⁰. In brief, spheres were dissociated into single cells and plated into 96-well plates in 0.2 ml of medium containing growth factors at increasing densities (1–100 cells per well) in triplicate. Cultures were left undisturbed for 14 days, and then the percentage of wells not containing spheres for each cell dilution was calculated and plotted against the number of cells per well. Linear regression lines

were plotted, and we estimated the minimal frequency of glioma cells endowed with stem cell capacity (the number of cells required to generate at least one sphere in every well = the stem cell frequency) based on the Poisson distribution and the intersection at the 37% level using Prism 6.0 software. Data represent the means of three independent experiments performed on different days.

The soft agar colony assay was performed by seeding human astrocytes at a density of 10,000 cells per well in 6-well plates in 0.3% agar in DMEM and 10% FBS. The number of colonies per well was determined using an Olympus 1X70 microscope equipped with a digital camera.

Subcutaneous xenograft glioma models. Mice were housed in a pathogen-free animal facility. All animal studies were approved by the IACUC at Columbia University (AAAQ2459; AAAL7600). Mice were 4–6 week old male and female athymic nude (*Nu/Nu*, Charles River Laboratories). No statistical method was used to pre-determine sample size. No method of randomization was used to allocate animals to experimental groups. Mice in the same cage were generally part of the same treatment. The investigators were not blinded during outcome assessment. In none of the experiments did tumours exceed the maximum volume allowed according to our IACUC protocol, specifically, 20 mm in the maximum diameter. 5×10^5 F3–T3 human astrocytes transduced with a lentivirus expressing the shRNA sequence against *PPARGC1A* or *ESRRG* or the empty vector were injected subcutaneously in the right flank in 150 µl of saline solution (five mice per group). 0.5×10^5 F3–T3;shTrp53 mGSCs and HRAS(12V);shTrp53 mGSCs transduced with lentivirus expressing two independent shRNA sequences against *PPARGC1A* were injected subcutaneously in the right flank in 100 µl of saline solution (five mice per group). Treatment with tigecycline (10 mice) or vehicle (8 mice) was performed in mice injected with 1×10^5 F3–T3;shTrp53 mGSCs when tumours reached 150–270 mm³ (10 days after injection). Tigecycline was diluted in saline pH 7 and administered at dose of 50 mg kg^{−1} body weight by intraperitoneal injection b.i.d. Tumour diameters were measured daily with a caliper and tumour volumes estimated using the formula: $(width^2 \times length) / 2 = V$ (mm³). Mice were euthanized when tumour size reached the maximum diameter allowed by our IACUC protocol (20 mm in the maximum diameter) or when mice displayed body weight loss equal to or greater than 20% of total body mass, or showed signs of compromised health or distress.

Plasmids, cloning, and lentivirus production. cDNAs for *FGFR3*, *PIN4*, *PKM2*, *GOLGIN84*, *DLG3*, *C1ORF50* and *PPARGC1A* were amplified by PCR and cloned into the pLOC vector in frame with Flag or V5 tag. F3–T3, F3–T3^{K508M} and FLAG-tagged *PEX1* were cloned into a pLVX-puro vector (Clontech). To generate PIN4(Y122A), PIN4(Y122F), PKM2(Y105A), GOLGIN84(Y42A), DLG3(Y673A) and C1orf50(Y131A), site-directed mutagenesis was performed using the QuickChange Site-Directed mutagenesis kit (Agilent) and the resulting plasmids were sequence verified. Lentivirus was produced by co-transfection of the lentiviral vectors with pCMV-ΔR8.1 and pCMV-MD2.G plasmids into HEK293T cells as previously described^{7,16}. shRNA sequences are: *PIN4* shRNA: GTCAGACACATTCTATGTGAAGTCGAGTTCACATAGAATGTGTCTGAC; *PPARGC1A* Hs-shRNA1: GCAGAGTATGACGATGGTATTCTCGAGAA TACCATCGTCATACTCTGC; *PPARGC1A* Hs-shRNA2: CCGTT ATACCTGTGATGCTTCTCGAGAAAGCATACAGGTATAACGG; *Ppargc1a* Mm-shRNA1: CCAGAACAAGAACAACGGTTTCTCGAGAAACCGT TGTCTTCTGTCTGG; *Ppargc1a* Mm-shRNA2: CCCATTGAGAAACAGAC TATCTCGAGATAGTCTTGTCTCAAATGGG; *ESRRG* shRNA1: CAAACAA AGATCGACACATTGCTCGAGCAATGTGTCGATCTTTGTTT; *ESRRG* shRNA2: CATGAAGCGCTGCAGGATTATCTCGAGATAATCCTGC AGCGCTTCATG.

For the generation of *PPARGC1A*-knockout F3–T3 human astrocytes with CRISPR–Cas9, guide RNA (gRNA) sequences were designed to target the coding sequence of *PPARGC1A* as described (<http://crispr.mit.edu/>). We designed two gRNAs against exon 1 of the *PPARGC1A* gene and validated three clones for loss of PGC1α expression. sgRNA sequence 1 (GGCGTGGGACATGTGCAACC) and 2 (ACCAGGACTCTGAGTCTGTA) were inserted by linker cloning in the lentiviral vector pLCiG2^{53,54}. Human astrocytes expressing the empty vector or F3–T3 were infected either with pLCiG2 and control, pLCiG2 and *PPARGC1A* gRNA1 or pLCiG2 and *PPARGC1A* gRNA2. After 72 h of infection, cells were seeded in a 96-well plate at a density of 0.6 cell per well. Two weeks later, colonies were isolated and *PPARGC1A* deletion was analysed by RT–qPCR and western blot. For acute expression of F3–T3 in human astrocytes, cells were first transduced with pLOC expressing Flag–PIN4(WT) or the Y122F mutant. Subsequently, cells were transduced with a pLKO.1-puro vector encoding shRNA targeting *PIN4*. The levels of endogenous and ectopically expressed proteins were then verified by immunoblotting. Finally, cells were transduced with a pLVX vector expressing F3–T3.

Generation of phospho-PIN4(Y122) antibody. The anti-phospho-PIN4 antibody was generated by immunizing rabbits with a short synthetic peptide

containing phosphorylated Y122 (underlined) (PVKTKFGYHIIMVE) (Yenzym Antibodies, LLC). A two-step purification process was applied. First, antiserum was cross-absorbed against the phospho-peptide matrix to purify antibodies that recognized the phosphorylated peptide. Next, the anti-serum was purified against the un-phosphorylated peptide matrix to remove non-specific antibodies. Antibodies were validated using lysates from cells transfected with PIN4(WT) or the phospho-mutant PIN4(Y122A).

Immunoblot and immunoprecipitation. For western blot, cells were lysed in RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP40, 0.5% sodium dodecyl sulphate, 1.5 mM Na_3VO_4 , 10 mM sodium fluoride, 10 mM sodium pyrophosphate, 10 mM β -glycerolphosphate and EDTA-free protease inhibitor cocktail, Roche). Lysates were cleared by centrifugation at 15,000 r.p.m. for 15 min at 4°C. Phospho-tyrosine immunoprecipitation was performed on cells that were freshly collected in cold PBS containing Na_3VO_4 and lysed in RIPA buffer. Subsequently, 800 μg of protein extract was incubated with 30 μl of phospho-tyrosine sepharose beads (P-Tyr-100, Cell Signaling Technology, 9419) in a final volume of 800 μl overnight at 4°C. Beads were washed five times with cold RIPA buffer and eluted by 2 \times SDS sample buffer. Immunoprecipitates were separated by SDS-PAGE and transferred to a nitrocellulose membrane. Membranes were blocked in TBS with 5% non-fat milk and 0.1% Tween-20, and probed with primary antibodies overnight at 4°C. PIN4 and Flag-PEX1 immunoprecipitation was performed on cells that were freshly collected in cold PBS and lysed in 50 mM Tris pH 8.0, 150 mM NaCl, 0.5% NP40, 1 mM EDTA, 10% glycerol, protease and phosphatase inhibitors. Subsequently, 2,000 μg of protein extract was incubated with PIN4 antibody (Abcam, ab155283) at a concentration of 0.6 μg mg^{-1} cell lysate or Flag-M2 agarose beads in a final volume of 1,000 μl overnight at 4°C. For PIN4 immunoprecipitation, Protein A/G Plus agarose beads (Santa Cruz Biotechnology) were added for 2 h at 4°C. Beads were washed five times with cold lysis buffer including 300 mM NaCl, and immunocomplexes were eluted with PIN4 or Flag-M2 peptide at room temperature for 4 h or 45 min, respectively. Immunoprecipitates were separated by SDS-PAGE and transferred to a nitrocellulose membrane. Membranes were blocked in TBS with 5% non-fat milk and 0.1% Tween-20, and probed with primary antibodies overnight at 4°C.

Antibodies and concentrations were: FGFR3 1:1,000 (Santa Cruz, B-9, sc-13121), PIN4 1:1,000 (Abcam, ab155283), PKM2 1:1,000 (Cell Signaling, 3198), DLG3 (also known as SAP102) 1:1,000 (Cell Signaling, 3733), GOLGIN84 1:2,000 (Santa Cruz, H-283, sc-134704), C1orf50 1:1,000 (Novus Biologicals, NBP1-81053), HGS 1:1,000 (Abcam, ab72053), FAK 1:1,000 (Cell Signaling, 3285), Paxillin 1:1,000 (BD Transduction, 610051), PGC1 α 1:500 (Santa Cruz, H300, sc-13067), PGC1 α 1:1,000 (Novus Biological, NBP104676), ERR γ 1:500 (Abcam, ab128930), ERR γ 1:500 (R7D, PP-H6812000), phospho-FRS2 1:1,000 (Cell Signaling, 3861), FRS2 1:1,000 (Santa Cruz, sc-8318), phospho-STAT3 1:1,000 (Cell Signaling, 9131), STAT3 1:1,000 (Santa Cruz, C-20, sc-482), phospho-AKT 1:1,000 (Cell Signaling, 4060), AKT 1:1,000 (Cell Signaling, 9272), phospho-ERK1/2 1:1,000 (Cell Signaling, 4370), ERK1/2 1:1,000 (Cell Signaling, 9102), β -actin 1:2,000 (Sigma, A5441), PEX1 1:500 (BD Biosciences, 611719), PEX6 1:500 (Stress Marq, SMC-470), NUP214 1:500 (Abcam, ab70497), SEC16A 1:500 (Abcam, ab70722), DHX30 1:500 (Novus Biologicals, NBP1-26203), SUN-2 1:500 (Abcam, ab124916), Flag 1:1,000 (Abcam, ab1162), retinoblastoma 1:1,000 (BD Pharmingen, 554136), α -tubulin 1:2,000 (Sigma, T5168), total OXPHOS 1:1,000 (Abcam, ab110411), MTCO1 1:1,000 (Abcam, ab14705). Secondary horseradish-peroxidase-conjugated antibodies were purchased from Pierce and Enhanced Chemiluminescence (Amersham) or Super Signal West Femto (Thermo Scientific) was used for detection.

RT-qPCR. Total RNA was prepared using the Trizol reagent (Invitrogen) and cDNA was synthesized using SuperScript II Reverse Transcriptase (Invitrogen) as described^{16,55}. RT-qPCR was performed with a Roche 480 thermal cycler, using SYBR Green PCR Master Mix (Applied Biosystems). RT-qPCR results were analysed by the $\Delta\Delta C_t$ method⁵⁶ using 18S or *Actb* as the housekeeping gene.

Human primers used for RT-qPCR were as follows. *UQCRC1* forward 5'-CACC GTGATGATGCTCTACC-3' and reverse 5'-CCACCACCATAAGTGCAGTC-3'; *POLRMT* forward 5'-TATTCATGGTGAAGGATGCC-3' and reverse 5'-TCTGT TCCACACCTTTTCG-3'; *NDUF4* forward 5'-TGCTTCAGTACACGA TCCC-3' and reverse 5'-CACACAGAGCTCCCATGAGT-3'; *MRPL15* forward 5'-TGCTTCCACCAGAAGAAGT-3' and reverse 5'-ACTTCTGCGCA GTTCAAGT-3'; *MCL1* forward 5'-GCATCGAACCATTAGCAGAA-3' and reverse 5'-TGCCACCTTCTAGGTCCTCT-3'; *MRPS30* forward 5'-TATTC CTCGTGGTTCATCGAA-3' and reverse 5'-CTCTGCGAGTTGCTTGGATA-3'; *TIMM10* forward 5'-CCTGGACCGATGTGTCTCTA-3' and reverse 5'-GCACCCT CTTCATCAGCTCT-3'; *NRF1* forward 5'-GGAAACGGCCTCATGTATTT-3' and reverse 5'-TCATCTAACGTGGCTCGAAG-3'; *ATP5G3* forward 5'-CCCAGAATG GTGTGTCTCAG-3' and reverse 5'-TTCCAATACCAGCACCAGAA-3'; *ABCE1* forward 5'-TCATTGATCAAGAGGTGCAGA-3' and reverse 5'-TAGACATC

AGCAGGTTTGCC-3'; *TIMM23* forward 5'-GGATTGAAGGAAACCAGAA-3' and reverse 5'-CCCTTGCTAGTCACCATATT-3'; *TFAM* forward 5'-GCTC AGAACCAGATGCAA-3' and reverse 5'-CACTCCGCCCTATAAGATC-3'; *TIMM9* forward 5'-TGAAGAGACCACCTGTTTCA-3' and reverse 5'-AAGGAGTCTGCTTTGGCT-3'; *TIMM44* forward 5'-TCCAA GACAGAGATGTCGGA-3' and reverse 5'-GATGTCGTTCTCGCACTGTT-3'; *VDAC1* forward 5'-AATGTGAATGACGGGACAGA-3' and reverse 5'-ACACGGG TCTCCAACTTCTT-3'; *NDUFA9* forward 5'-TATGCATCGGTTTGGTCTTA-3' and reverse 5'-GACCAACGAAAGCAAAGGAT-3'; *NDUFV2* forward 5'-AAAG GCAGAATGGGTGGTT-3' and reverse 5'-CTTTCCAACTGGCTTTTCGAT-3'; *COX5A* forward 5'-GATGCTCGCTGGGTAACATA-3' and reverse 5'-GGGCTCTG GAACATATCAT-3'; *NDUFS5* forward 5'-TGGTGAACAGCCCTACAAGA-3' and reverse 5'-TCTGCCGAGTATAACCGAT-3'; *NDUFA2* forward 5'-CGTCAG GGACTTCATTGAGA-3' and reverse 5'-AAGGGACATTCTGCTCTTGG-3'; *PITRM1* forward 5'-TCTCGGATGAGATGAAGCAG-3' and reverse 5'-CCCAG TGCCGAGGTATCTAT-3'; *CISD1* forward 5'-TGACTTCCAGTTCCAGCGTA-3' and reverse 5'-GATAACCAATTGCAGCTGTCC-3'; *ATP5G1* forward 5'-AGCTCT GATCCGCTGTGTGA-3' and reverse 5'-GGAAGTTGCTGTAGGAAGGC-3'; *MRPL12* forward 5'-TCAACGAGCTCCTGAAGAAA-3' and reverse 5'-GTGTC CGTTCTTTCGCTATG-3'; *IDH3A* forward 5'-CCGACCATGTGTCTCTATCG-3' and reverse 5'-GCACGACTCCATCAACAATC-3'; *18S* forward 5'-CGCCG CTAGAGGTGAATTC-3' and reverse 5'-CTTTTCGCTCTGGTCCGCTCTT-3'; *PPARGC1A* forward 5'-CTCACACCAAAACCCACAGAG-3' and reverse 5'-GTGT TGTGACTGCGACTGTG-3'; *PEX1* forward 5'-AGTACCAGCCTGCA TTCTT-3' and reverse 5'-ATGGGAACATGGCTTGAGAA-3'.

Mouse primers used in RT-qPCR were as follows. *Ppargc1a* forward 5'-GACAG CTTTCTGGGTGGATT-3' and reverse 5'-CGCAGGCTCATTGTTGTACT-3'; *Actb* forward 5'-GATGACGATATCGCTGCGCTG-3' and reverse 5'-GTACGACCAGAGGCATACAGG-3'.

Quantification of mitochondrial DNA content. Total DNA was isolated using the Puregene Blood Core kit (Qiagen) according to the manufacturer's instructions. Mitochondrial DNA content was measured using real-time quantitative PCR (qPCR) as previously described⁵⁷. In brief, relative quantification of mitochondrial DNA content for each sample was determined using a set of mitochondrial specific primers: mt-Mito: forward 5'-CACTTCCACACAGACATCA-3', reverse 5'-TGGTTAGGCTGGTGTAGGG-3'; and a set of nuclear-specific primers: *B2M* forward 5'-TGTTCTGCTGGGTAGCTCT-3' and reverse 5'-CCTCCATGATG CTGCTTACA-3'. qPCR was performed with a Roche 480 thermal cycler, using SYBR Green PCR Master Mix (Applied Biosystems). qPCR conditions used were: 95°C for 10 min followed by 40 cycles at 95°C for 15 s, 72°C for 15 s followed by a melting cycle going up to 95°C. Primer specificity was determined by melt curve analysis and agarose gel electrophoresis, confirming a single band of the amplification product. The relative mitochondrial DNA content was calculated using the $2^{-\Delta\Delta C_t}$ method where ΔC_t is $C_t^{\text{mt-Mito}} - C_t^{\text{B2M}}$.

Mitochondria analysis by flow cytometry. Cells were seeded in 60-mm dishes and cultured in DMEM containing 10% FBS. Cells were washed once with FBS-free medium. Mitotracker Red (Life Technologies, M7212) was added to a final concentration of 20–40 nM and incubated for 20–30 min at 37°C. Cells were then quickly washed with PBS, trypsinized, collected in phenol-red-free medium and incubated for 10 min at 37°C in the dark before analysis. Unstained cells were used as a negative control. Acquisition was performed on LSR II Flow Cytometer (BD Biosciences) on the basis of forward and sideward scatter parameters and Texas red fluorescence using BD FACSDiva software. Eight to ten thousand events from each sample were evaluated. Data were analysed using the FCS Express 6 Flow software (De novo Software).

Metabolic assays. **Measurement of oxygen consumption rate and extracellular acidification rate.** The functional status of mitochondria in cells expressing F3–T3 was determined by analysing multiple parameters of oxidative metabolism using the XF96 Extracellular Flux Analyzer (Agilent), which measures the extracellular flux changes of oxygen and protons. Cells were plated in XF96-well microplates (6,000–7,000 cells per well) in a final volume of 80 μl of DMEM medium (25 mM glucose, 2 mM glutamine) supplemented with 10% FBS, 48 h before the assay. For experiments requiring AZD4547 treatment, cells were plated as previously described in the presence of 150 nM of AZD4547. For the mitochondrial stress test, cells were washed twice with 200 μl of XF Assay Medium Modified DMEM (Agilent), supplemented with 25 mM glucose, 2 mM glutamine (XF-Mito-MEM) and incubated at 37°C in the absence of CO_2 for 1 h before the assay in 180 μl per well of XF-Mito-MEM. The ports of the sensor cartridge were sequentially loaded with 20 μl per well of the appropriate compound: the ATP coupler oligomycin (Sigma, O4876), the uncoupling agent carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone (FCCP, Sigma C2920) and the complex I inhibitor rotenone (Sigma, R8875). Compound concentration used for the different cell lines are indicated in Supplementary Table 9a. For the glycolysis stress test, cells were washed twice with

200 µl of XF Assay Medium Modified DMEM supplemented with 2 mM glutamine (XF-Glyco-MEM) and incubated at 37 °C in the absence of CO₂ for 1 h before the assay in 180 µl per well of XF-Mito-MEM. The ports of the sensor cartridge were sequentially loaded with 20 µl per well of the appropriate compound: glucose (10 mM final concentration; Sigma, 8769), oligomycin (2 µM final concentration; Sigma, O4876) and 2-DG (100 mM final concentration; Sigma, 8375). OCR and ECAR were measured through 16 rates: 4 rates under basal conditions, 4 rates after oligomycin or glucose injection, 4 rates after FCCP or oligomycin injection, and 4 rates after rotenone or 2-DG injection, for OCR or ECAR evaluation, respectively. The protocol was: mix (2 min), wait (1 min) and measure (2 min). OCR and ECAR values were normalized to the number of cells per well. The ratio OCR:ECAR was determined by dividing the normalized value of rate 4 of the OCR (basal condition) with the normalized value of rate 8 of the ECAR (after glucose injection).

ATP assay. Cells were cultured in DMEM medium containing 25 mM glucose, 2 mM glutamine and 10% FBS. After 24 h, 5,000 cells per 130 µl were plated in opaque white 96-well plates in 5 mM glucose, 2 mM glutamine and 0.2% FBS DMEM medium. CellTiterGlo assay reagent (Promega, G7570) was added 12–16 h later, according to the manufacturer's instructions and luminescence was measured using a GloMax-Multi+ Microplate Multimode Reader (Promega). For experiments testing the effect of oligomycin, cells were cultured in glucose-free DMEM medium for 24 h and ATP levels were determined as described above.

Cell growth assay. Time-course analysis of cellular growth of human astrocytes expressing F3–T3 or the empty vector was performed by plating 12,500 cells per well in triplicate in 6-well plate in DMEM containing 10% FBS. After 24 h, cells were washed and cultured in glucose-free DMEM medium containing 0.2% FBS and supplemented with 25 mM glucose or 25 mM galactose in the presence or absence of 100 nM oligomycin. Viable cells were scored every two days by Trypan blue exclusion. Survival assays of human and mouse GSCs treated with mitochondrial inhibitors were performed by plating 25,000 cells per well in 12-well plates in triplicate. Cells were counted after 72 h.

Gene set enrichment analysis (GSEA) for ROS detoxification genes. We generated a gene set of 46 genes participating in ROS detoxification programs by combining genes extracted from specific references^{20,58} and the detoxification of reactive oxygen species reactome pathway R-HSA-3299685 (<http://reactome.org/pages/download-data/>). The full list of genes is reported in Supplementary Table 7c. The GSEA analysis was performed comparing the gene expression of F3–T3-expressing and vector control-transduced human astrocytes using default settings⁵⁹.

Analysis of protein biosynthesis and ROS by high content microscopy. Human astrocytes transduced as indicated in the figure legends were plated at a density of 6,000 cells per well in 96-well clear-bottom black plates (Greiner) 18 h before the analysis in preparation for the evaluation of both protein biosynthesis and cellular ROS. Protein biosynthesis was detected by the Click-iT Plus OPP Alexa Fluor-594 protein synthesis assay kit (Molecular Probes, C10457). Cells were incubated in the dark with O-propargyl-puromycin (OPP) reagent at a concentration of 10 µM for 30 min. Identical samples were treated with CHX at a concentration of 30 µM for 30 min before addition of OPP reagent and used as negative controls. Samples were washed with Click-iT rinse buffer, fixed in 3.7% formaldehyde for 15 min followed by permeabilization in 0.5% Triton X-100 for 15 min. Click-iT OPP reaction cocktail was then added for 30 min followed by one wash in Click-iT rinse buffer and nuclear counterstaining with DAPI. Acquisition of fluorescence intensity was performed using an IN Cell Analyzer 2000 (GE Healthcare) equipped with a 2,048 × 2,048 CCD camera. Assay conditions were: two-colour assay (DAPI and Cy3), 20× objective, exposure time 0.5 ms for Cy3 and 0.1 ms for DAPI. Four fields around the centre of each well, including 2,000 to 6,000 cells, were imaged. Data were analysed using IN Cell Investigator software (GE Healthcare). Fluorescence intensity was normalized to the number of cells as determined by the number of DAPI-positive objects in each well.

For determination of cellular ROS, cells were incubated in the dark with CellROX Deep Red reagent at concentration of 2.5 µM for 30 min at 37 °C. Identical samples were treated with *N*-acetyl-L-cysteine at concentration of 5 µM for 2 h before addition of CellROX Deep Red reagent and used as negative controls. Samples were washed once with PBS, fixed in 3.7% formaldehyde for 15 min followed by two additional washes in PBS and nuclear counterstaining with DAPI. Acquisition of fluorescence intensity was performed as described above. Assay conditions were: two-colour assay (DAPI and Cy5), 20× objective, exposure time 0.8 ms for Cy5 and 0.1 ms for DAPI. Four fields around the centre of each well, including 2,000 to 6,000 cells, were imaged. Data were analysed using IN Cell Investigator software (GE Healthcare). Fluorescence intensity was normalized to the number of cells as determined by the number of DAPI-positive objects in each well.

Immunofluorescence of cultured cells and primary tissue. Cells were fixed with 4% paraformaldehyde containing 4% sucrose, permeabilized with 0.1–1% Triton

X-100, 0.1% BSA in TBS for 4 min at 4 °C, and blocked with 3% BSA, 0.05% Triton X-100 in TBS. The primary antibodies used were as follows: phospho-PIN4 (1:100); PMP70 (Sigma, SAB4200181, 1:150); PEX1 (BD Bioscience, 611719, 1:100); FGFR3 (Santa Cruz, sc-13121 B9, 1:1,000). Secondary antibodies were anti-mouse Alexa Fluor-647, anti-rabbit Alexa Fluor-568, or Cy3-conjugated (Molecular Probes, Invitrogen). Nuclei were stained with DAPI (Sigma).

Fluorescence microscopy was performed on a Nikon A1R MP microscope using a 100×, 1.45 Plan Apo Lambda lens. Images were recorded with a z-optical spacing of 0.15 µm and analysed using the NIS Elements Advanced Research software (Nikon Instruments). The number of peroxisomes per cell was scored as the average of PMP70⁺ in five z sections (one at the equatorial plane and two above and two below the equatorial plane). Quantification of PEX1 fluorescence intensity was performed on maximum intensity images of z sections. After calibration and thresholding, the integrated density (product of the area and the mean intensity value (IMFI)) was averaged between 30 cells in at least six representative pictures per sample.

Tissue preparation and immunostaining on mouse and human tissues were performed as previously described^{7,50,60}. The human GBM samples analysed by immunostaining had been stored in the Onconeurotek Tumourbank (certified NF S96 900), and received the authorization for analysis from ethical committee (CPP Ile de France VI, A39II), and French Ministry for research (AC 2013-1962). In brief, tumour sections were deparaffinized in xylene and rehydrated in a graded series of ethyl alcohol. Antigen retrieval was performed in citrate solution pH 6.0 using a decloaking chamber (10 min for phospho-PIN4 and 7 min for COXIV, VDAC, NDUFS4 and FGFR3). Primary antibodies were incubated at 4 °C overnight: COXIV (Cell Signaling, 4850, 1:1,500), VDAC1 (Abcam, ab14734, 1:700), NDUFS4 (Abcam, ab55540, 1:700), phospho-PIN4(Y122) (1:200) and FGFR3 (Santa Cruz, B9, sc-13121, 1:500). Sections were incubated in biotinylated secondary antibody for 1 h, followed by 30 min of streptavidin–HRP-conjugated (Vector Laboratories) for phospho-PIN4(Y122), FGFR3, VDAC1, and NDUFS4 or HRP-conjugated anti-rabbit secondary antibody (DAKO) for COXIV and TSA–Cy3 or TSA–Fluorescein (Perkin-Elmer). Nuclei were counterstained with DAPI (Sigma). Images were acquired using 20× magnification using an Olympus 1X70 microscope equipped with digital camera. Quantification of fluorescence intensity was performed using NIH ImageJ software. After calibrating and standardizing the 8-bit grayscale images, the integrated density (IMFI) was averaged between three 20× representative pictures per sample section.

Drosophila studies. The *UAS-F3–T3* flies were generated by inserting the human F3–T3 fusion gene into the pACU2 plasmid followed by embryo injection of the plasmid and selection of the correct transgenic fly. All other genotypes were established through standard genetics. *repo-Gal4* was used to drive gene expression in the glial lineage. *UAS-eGFP* or *UAS-mRFP* were introduced to visualize and quantify tumour volume. *repo-Gal4;UAS-dEGFR^Δ;UAS-Dp110^{CAAX}* (as previously described²⁴) and *repo-Gal4;UAS-F3–T3* stocks were balanced over the *CyoWeeP* and *TM6B* balancers. *srl* RNAi lines were obtained from the Bloomington *Drosophila* Stock Center (BDSC) and the Vienna *Drosophila* Resource Center (VDRC): *P{KK100201}/VIE-260B* (VDRC v103355), *y¹ sc^{*} v¹*; *P{TriP.GL01019}/attP40* (BDSC 57043), *y¹ sc^{*} v¹*; *P{TriP.HMS00857}/attP2* (BDSC 33914), and *y¹ sc^{*} v¹*; *P{TriP.HMS00858}/attP2* (BDSC 33915). The following *ERR* RNAi lines were used: *y¹ v¹*; *P{TriP.JF02431}/attP2* (BDSC 27085), *y¹ v¹*; *P{TriP.HMC03087}/attP2* (BDSC 50686) and *P{KK108422}/VIE-260B* (VDRC v108349). *y¹ v¹*; *P{UAS-GFP.VALIUM10}/attP2* (BDSC 35786) was used as a control for RNAi experiments.

Fly culture, immunohistochemistry and imaging. Flies were mated and maintained at 29 °C. Fly larvae were retrieved at late third instar stage for brain dissections followed by fixation and immunohistochemical analysis. Larval brains were dissected, fixed and stained as previously described⁶¹. In brief, third instar larval brains were dissected in PBS, fixed in 4% paraformaldehyde solution for 20 min at room temperature, and incubated with primary antibodies, including: rat anti-phospho-histone H3 (Abcam, ab10543, 1:300) and mouse anti-Repo (Developmental Study Hybridoma Bank, 1:60) overnight at 4 °C and secondary antibody for 2 h at room temperature. Fluorescence images were acquired using a Leica SP8 confocal microscope.

Image analysis. To determine tumour volume, we acquired image stacks using a Leica SP8 confocal microscope with a z-step size of 5.0 µm per optical slice using a 20× objective throughout the entire thickness of the brain and ventral nerve cord. The confocal LIF files were converted into Imaris files using ImarisFileConverter 6.4.2. All subsequent image processing was conducted with Imaris 5.5 software. z-series stacks were used to make three-dimensional reconstructions. A smooth level of 1.0 was used on every measurement for consistency. Brain tumour volumes were quantified using three-dimensional reconstructions.

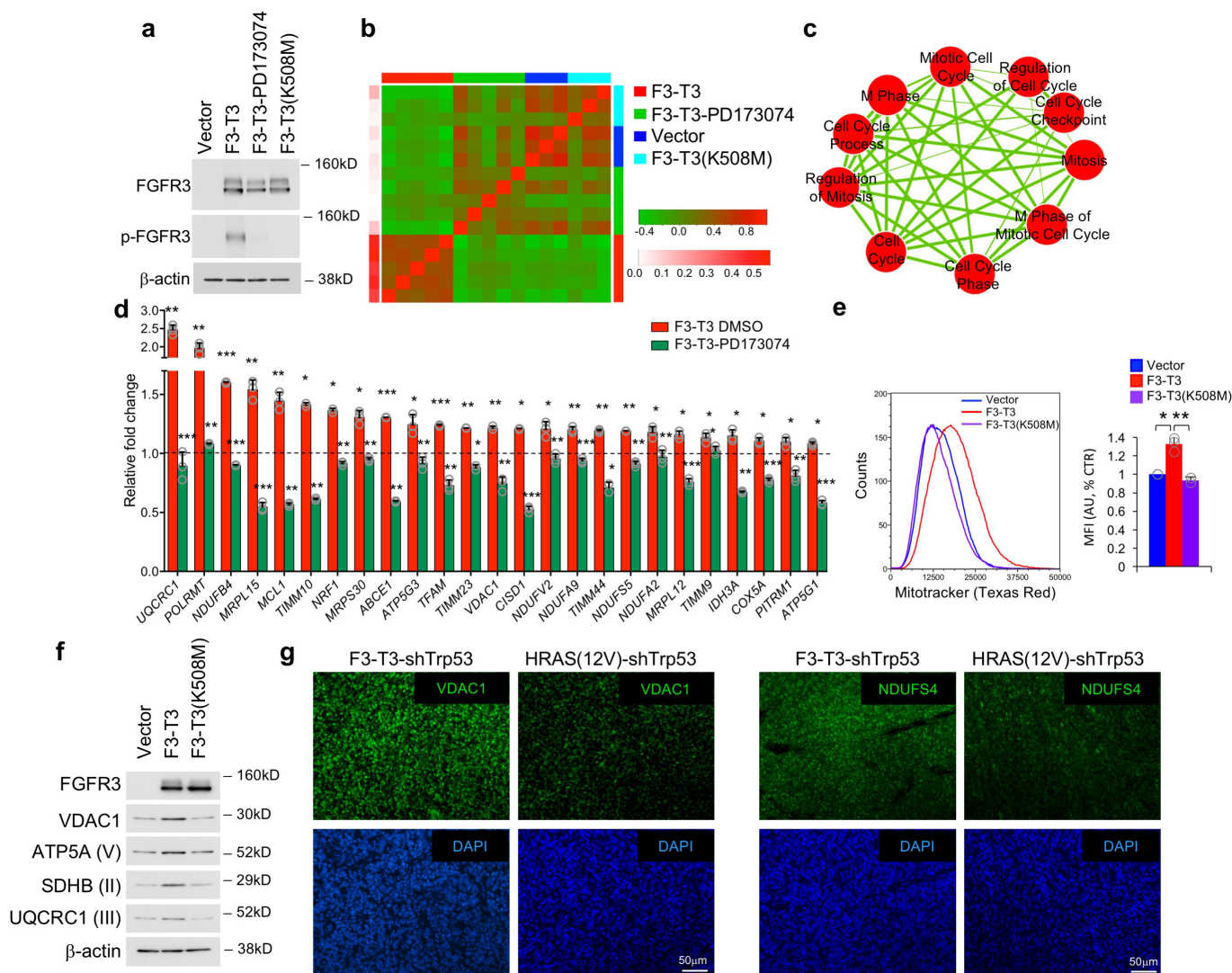
Statistical analysis. In general, two to four independent experiments were performed. Comparisons between groups were analysed by *t*-test with Welch correction (two-tailed, unequal variance) and/or the MWW non-parametric test when

appropriate. Results in graphs are expressed as mean \pm s.d. or mean \pm s.e.m. for the indicated number of observations. All statistical analyses were performed and *P* values were obtained using the GraphPad Prism software 6.0 or the R software (<https://www.r-project.org>) and are reported in Source Data.

Code availability. A collection of the R procedures to perform ee-MWW is available at <http://github.com/miccec/yaGST>. The RGBM package is available from CRAN at <https://cran.r-project.org/web/packages/RGBM/index.html>.

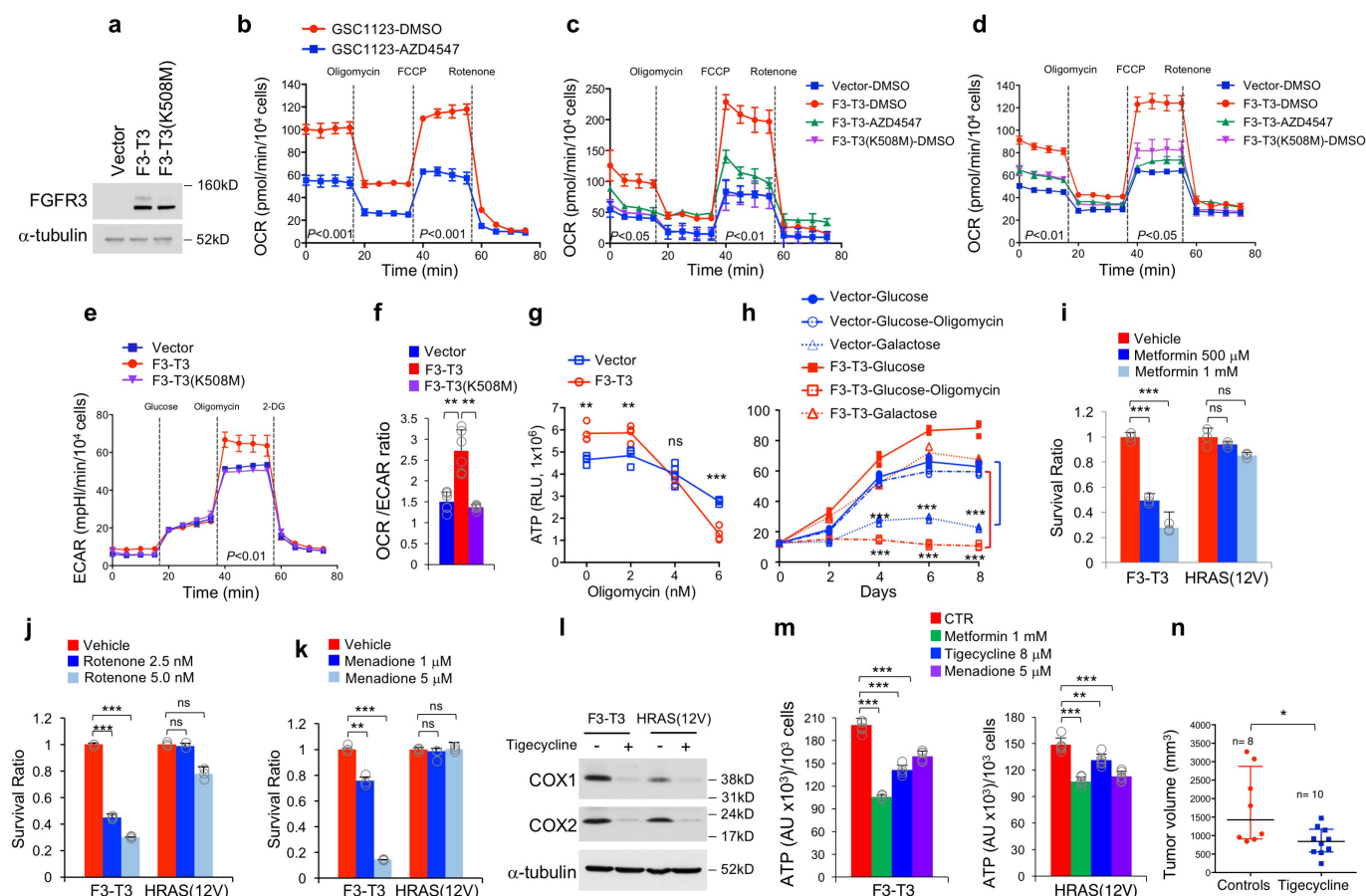
Data availability. Transcriptomic microarray gene expression data have been deposited in ArrayExpress with accession number E-MTAB-6037. Source data for western blot are provided in Supplementary Fig. 1. Source Data for Figs 1–4 and Extended Data Figs 1, 2, 4, 7–10 are included in the online version of the paper.

28. Isserlin, R., Merico, D., Voisin, V. & Bader, G. D. Enrichment Map — a Cytoscape app to visualize and explore OMICS pathway enrichment results. *F1000Res.* **3**, 141 (2014).
29. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431–432 (2011).
30. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE* **5**, e13984 (2010).
31. Stransky, N., Cerami, E., Schalm, S., Kim, J. L. & Lengauer, C. The landscape of kinase fusions in cancer. *Nat. Commun.* **5**, 4846 (2014).
32. Yoshihara, K. *et al.* The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene* **34**, 4845–4854 (2015).
33. Iyer, M. K., Chinnaiyan, A. M. & Maher, C. A. ChimeraScan: a tool for identifying chimeric transcription in sequencing data. *Bioinformatics* **27**, 2903–2904 (2011).
34. Risso, D., Schwartz, K., Sherlock, G. & Dudoit, S. GC-content normalization for RNA-seq data. *BMC Bioinformatics* **12**, 480 (2011).
35. Ceccarelli, M. *et al.* Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell* **164**, 550–563 (2016).
36. Lum, P. Y. *et al.* Extracting insights from the shape of complex data using topology. *Sci. Rep.* **3**, 1236 (2013).
37. Roweis, S. T. & Saul, L. K. Nonlinear dimensionality reduction by locally linear embedding. *Science* **290**, 2323–2326 (2000).
38. Nielson, J. L. *et al.* Topological data analysis for discovery in preclinical spinal cord injury and traumatic brain injury. *Nat. Commun.* **6**, 8581 (2015).
39. Singh, G., Memoli, F. & Carlsson, G. E. Topological methods for the analysis of high dimensional data sets and 3D object recognition. In *SPBG* (eds Botsch, M. & Pajarola, R.) 91–100 (Citeseer, 2007).
40. Rizvi, A. H. *et al.* Single-cell topological RNA-seq analysis reveals insights into cellular differentiation and development. *Nat. Biotechnol.* **35**, 551–560 (2017).
41. Rolland, D. *et al.* Global phosphoproteomic profiling reveals distinct signatures in B-cell non-Hodgkin lymphomas. *Am. J. Pathol.* **184**, 1331–1342 (2014).
42. Kessner, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534–2536 (2008).
43. Fenyő, D. & Beavis, R. C. A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes. *Anal. Chem.* **75**, 768–774 (2003).
44. MacLean, B., Eng, J. K., Beavis, R. C. & McIntosh, M. General framework for developing and evaluating database scoring algorithms using the TANDEM search engine. *Bioinformatics* **22**, 2830–2832 (2006).
45. Keller, A., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392 (2002).
46. Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646–4658 (2003).
47. Pedrioli, P. G. Trans-proteomic pipeline: a pipeline for proteomic analysis. *Methods Mol. Biol.* **604**, 213–238 (2010).
48. Fermin, D., Basrur, V., Yocum, A. K. & Nesvizhskii, A. I. Abacus: a computational tool for extracting and pre-processing spectral count data for label-free quantitative proteomic analysis. *Proteomics* **11**, 1340–1345 (2011).
49. Beausoleil, S. A., Villén, J., Gerber, S. A., Rush, J. & Gygi, S. P. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* **24**, 1285–1292 (2006).
50. Lee, S. B. *et al.* An ID2-dependent mechanism for VHL inactivation in cancer. *Nature* **529**, 172–177 (2016).
51. Sonoda, Y. *et al.* Formation of intracranial tumors by genetically modified human astrocytes defines four pathways critical in the development of human anaplastic astrocytoma. *Cancer Res.* **61**, 4956–4960 (2001).
52. Rutka, J. T. *et al.* Establishment and characterization of five cell lines derived from human malignant gliomas. *Acta Neuropathol.* **75**, 92–103 (1987).
53. Malina, A. *et al.* Adapting CRISPR/Cas9 for functional genomics screens. *Methods Enzymol.* **546**, 193–213 (2014).
54. Ran, F. A. *et al.* Genome engineering using the CRISPR–Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
55. Zhao, X. *et al.* The HECT-domain ubiquitin ligase Huwe1 controls neural differentiation and proliferation by destabilizing the N-Myc oncoprotein. *Nat. Cell Biol.* **10**, 643–653 (2008).
56. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402–408 (2001).
57. Ajaz, S., Czajka, A. & Malik, A. Accurate measurement of circulating mitochondrial DNA content from human blood samples using real-time quantitative PCR. *Methods Mol. Biol.* **1264**, 117–131 (2015).
58. Vazquez, F. *et al.* PGC1 α expression defines a subset of human melanoma tumors with increased mitochondrial capacity and resistance to oxidative stress. *Cancer Cell* **23**, 287–301 (2013).
59. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
60. Lasorella, A., Rothschild, G., Yokota, Y., Russell, R. G. & Iavarone, A. Id2 mediates tumor initiation, proliferation, and angiogenesis in Rb mutant mice. *Mol. Cell. Biol.* **25**, 3563–3574 (2005).
61. Zhu, S. *et al.* The bHLH repressor Deadpan regulates the self-renewal and specification of *Drosophila* larval neural stem cells independently of Notch. *PLoS ONE* **7**, e46724 (2012).



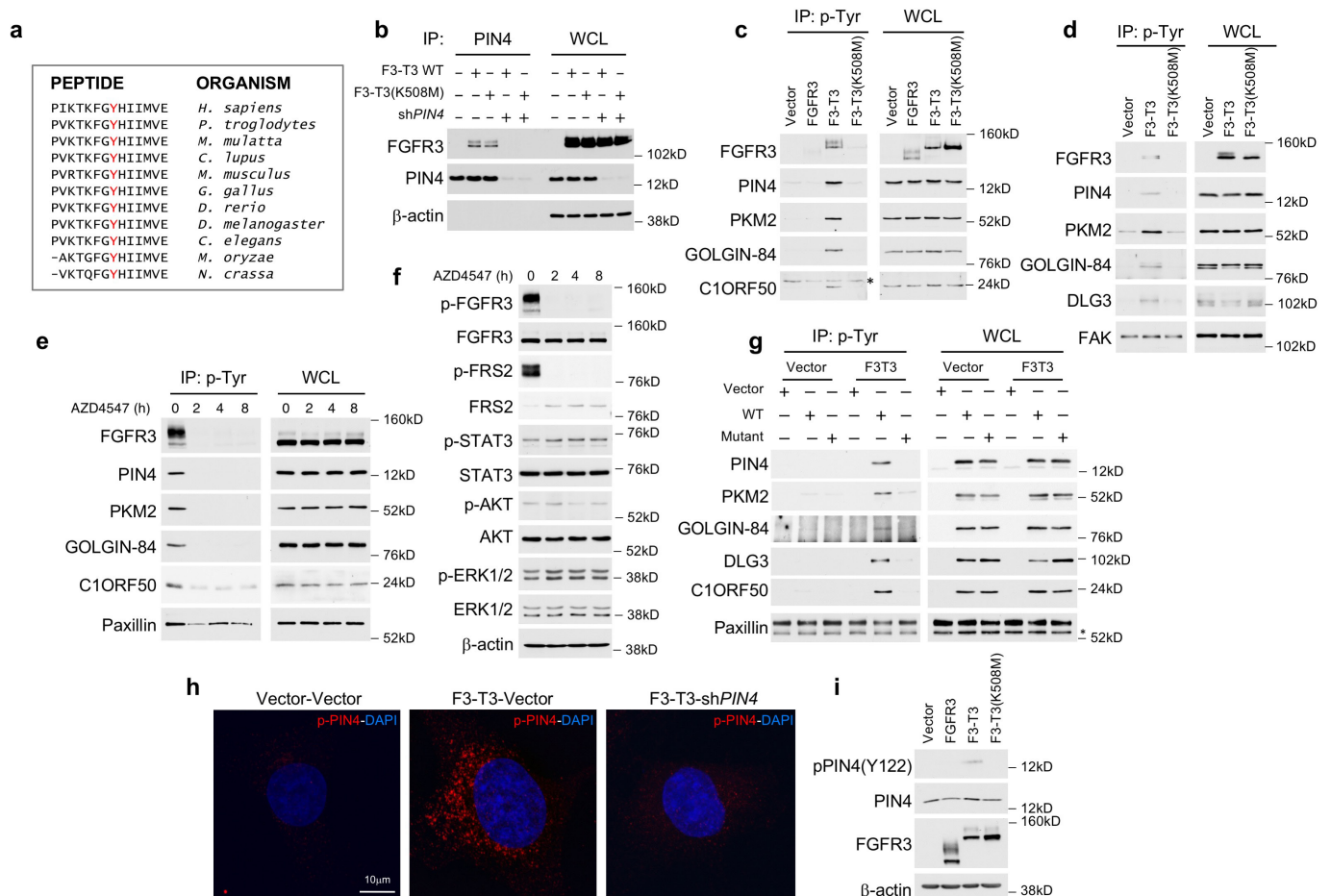
Extended Data Figure 1 | Activation of mitosis and mitochondria by F3-T3. **a**, Immunoblot analysis of FGFR3 and phospho-FGFR3 in F3-T3 human astrocytes treated with DMSO or PD173074, or human astrocytes expressing F3-T3(K508M) or vector. β-Actin is shown as a loading control. Experiment was repeated at least five times with similar results. **b**, Heat map of correlations among F3-T3 human astrocytes, F3-T3 human astrocytes treated with PD173074, and human astrocytes expressing vector or F3-T3(K508M). Top and right track colours represent sample type; left track colour scale represents correlation between each sample and the F3-T3 group. F3-T3 human astrocytes and F3-T3 human astrocytes treated with PD173074 ($n = 5$ biologically independent samples per group). Human astrocytes expressing vector or F3-T3(K508M) ($n = 3$ biologically independent samples per group). **c**, Enrichment map network of GO categories scoring as significant ($Q < 10^{-6}$ in each comparison) from three independent GSEAs (F3-T3 human astrocytes versus F3-T3 human astrocytes treated with PD173074; F3-T3- versus F3-T3(K508M)-expressing human astrocytes; F3-T3- versus vector-expressing human astrocytes). Nodes represent GO terms and lines indicate their connectivity. Size of nodes is proportional to enrichment significance and thickness of lines indicates the fraction of genes shared between the

groups. **d**, RT-qPCR of vector- or F3-T3-expressing human astrocytes treated with vehicle (DMSO) or PD173074 for 12 h. Data are fold change relative to vector (dotted line) of one representative experiment out of two independent experiments (data are mean \pm s.d., $n = 3$ technical replicates). P values were calculated using a two-tailed t -test with unequal variance; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. For a complete list of P values see Source Data. **e**, Left, analysis of mitochondrial mass by MitoTracker FACS analysis in human astrocytes expressing F3-T3, F3-T3(K508M) or vector. Right, quantification of mean fluorescence intensity (MFI). Data are mean \pm s.d. of three (vector and F3-T3) and two (F3-T3(K508M)) independent experiments. $*P < 0.05$, $**P < 0.01$; two-tailed t -test with unequal variance. **f**, Immunoblot analysis of mitochondrial proteins in human astrocytes expressing F3-T3, F3-T3(K508M) and vector. Experiment was repeated independently three times with similar results. **g**, Representative micrographs of VDAC1 and NDUFS4 immunofluorescence (top, green) in F3-T3;shTrp53 and HRAS(12V);shTrp53 mGSCs. DAPI staining of nuclei is shown as an indication of cellular density (bottom, blue). Experiment was repeated independently twice with similar results. Molecular weights are indicated on all immunoblots.



Extended Data Figure 2 | F3-T3 induces sensitivity to inhibitors of mitochondrial metabolism. **a**, Immunoblot analysis using the FGFR3 antibody in human astrocytes expressing vector, F3-T3 or F3-T3(K508M). α -Tubulin is shown as a loading control. Experiment was repeated five times with similar results. **b**, OCR of GSC1123 cells expressing F3-T3 in the presence or absence of AZD4547. Data are mean \pm s.d. ($n = 6$ technical replicates) of one representative experiment out of two independent experiments. $P < 0.001$ for rate 1-4 and 9-12, two-tailed t -test with unequal variance. **c**, OCR of RPE cells expressing F3-T3, F3-T3(K508M) or the empty vector in the presence or absence of AZD4547. Data are mean \pm s.d. ($n = 3$ technical replicates) of one representative experiment out of three independent experiments performed in triplicate with similar results. $P < 0.05$ for rate 1-4; $P < 0.001$ for rate 9-12; two-tailed t -test with unequal variance. **d**, OCR of U251 cells expressing F3-T3, F3-T3(K508M) or the empty vector in the presence or absence of AZD4547. Data are mean \pm s.d. ($n = 3$ technical replicates) of one representative experiment out of two independent experiments performed in triplicate with similar results. $P < 0.01$ for rate 1-4; $P < 0.001$ for rate 9-12; two-tailed t -test with unequal variance. **e**, ECAR of human astrocytes expressing F3-T3, F3-T3(K508M) or the empty vector. Data are mean \pm s.d. ($n = 3$ technical replicates) of one representative experiment out of two independent experiments performed in triplicate with similar results. $P < 0.01$ for rate 9-12; two-tailed t -test with unequal variance. **f**, Ratio between OCR (rate 4) and ECAR (rate 8) in human astrocytes expressing F3-T3, F3-T3(K508M) or vector. Data are mean \pm s.d. ($n = 6$ replicates) of two independent experiments each performed in triplicate. $P < 0.01$; two-tailed t -test with unequal variance. **g**, Quantification of ATP production in human astrocytes expressing F3-T3 or vector following treatment with the indicated concentrations of oligomycin for 72 h. Data are independent

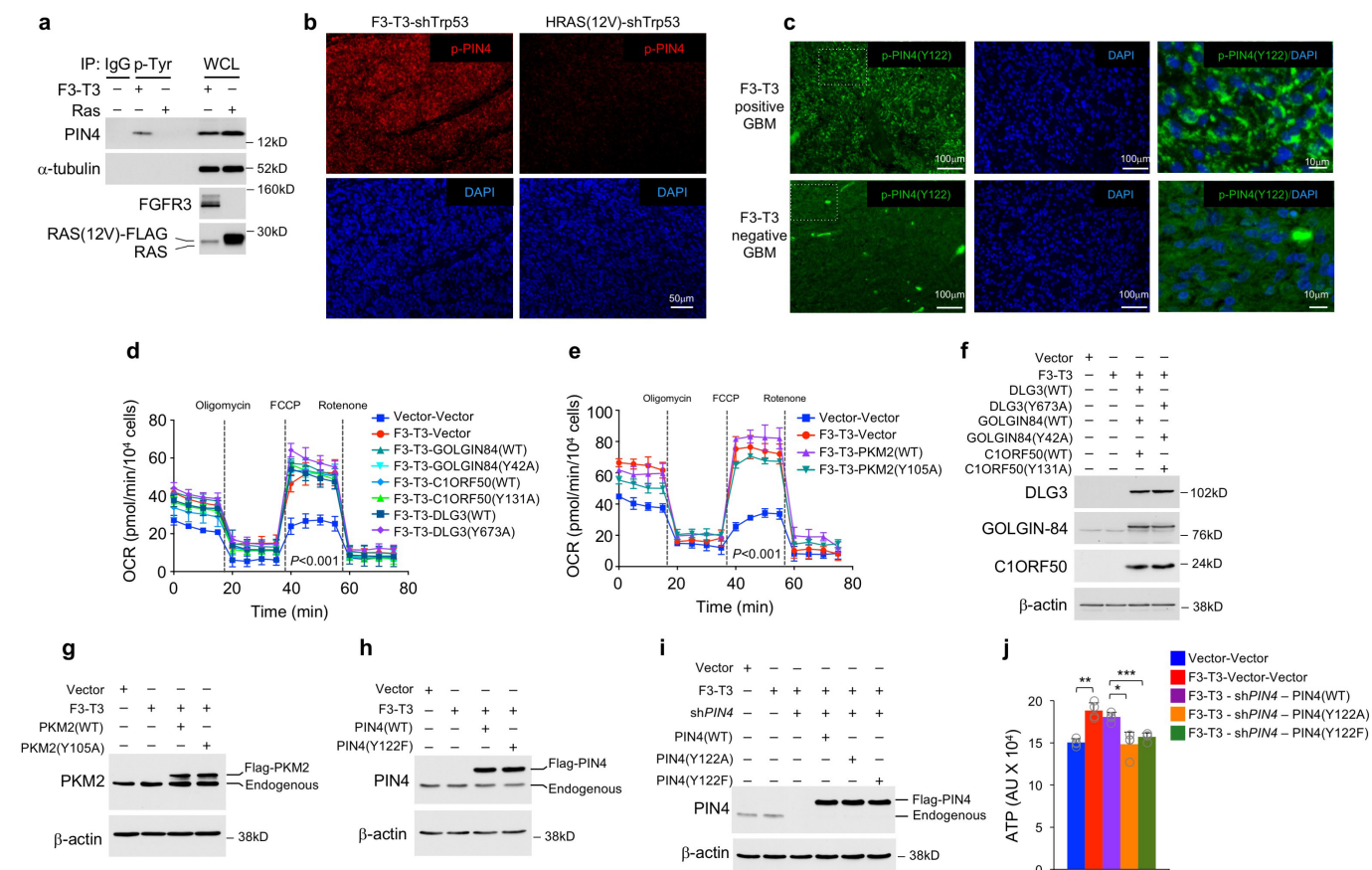
technical replicates ($n = 4$) and means (connecting lines) of one representative experiment out of two independent experiments performed with similar results. $**P < 0.01$; $***P < 0.001$; two-tailed t -test with unequal variance. **h**, Time-course analysis of cellular growth of human astrocytes expressing F3-T3 or vector cultured in the presence of glucose (25 mM) or galactose (25 mM) with or without oligomycin (100 nM). Data are independent technical replicates ($n = 3$) of one representative experiment out of two independent experiments performed with similar results. $***P < 0.001$; two-tailed t -test with unequal variance. **i-k**, Survival ratio of F3-T3;shTrp53 and HRAS(12V);shTrp53 mGSCs treated for 72 h with vehicle or metformin (**i**), rotenone (**j**) or menadione (**k**) at the indicated concentrations. Data are mean \pm s.d. ($n = 3$ technical replicates) of one representative experiment out of two independent experiments performed with similar results. $**P < 0.01$; $***P < 0.001$; two-tailed t -test with unequal variance. **l**, Western blot analysis of COX1 and COX2 proteins in F3-T3;shTrp53 and HRAS(12V);shTrp53 mGSCs treated with vehicle or tigecycline at a concentration of 8 μ M for 72 h. α -Tubulin is shown as a loading control. Experiment was independently repeated twice with similar results. **m**, Quantification of cellular ATP in F3-T3;shTrp53 (left) and HRAS(12V);shTrp53 (right) mGSCs treated with vehicle or metformin (1 mM), tigecycline (8 μ M) or menadione (5 μ M) for 16 h. Data are mean \pm s.d. of one experiment ($n = 6$ technical replicates). $**P < 0.01$; $***P < 0.001$; two-tailed t -test with unequal variance. **n**, Quantification of tumour volume of F3-T3;shTrp53 mGSCs in control and tigecycline-treated mice. Data are tumour volumes (median with interquartile range) at day 6 of treatment, a time when all mice were still in the study; $n = 8$ for control (median = 1,427 mm³) and $n = 10$ for tigecycline-treated mice (median = 843.4 mm³). $*P < 0.05$; two-sided Mann-Whitney U -test. Molecular weights are indicated in immunoblots.



Extended Data Figure 3 | Phosphorylation of Y122 of PIN4 by F3-T3.

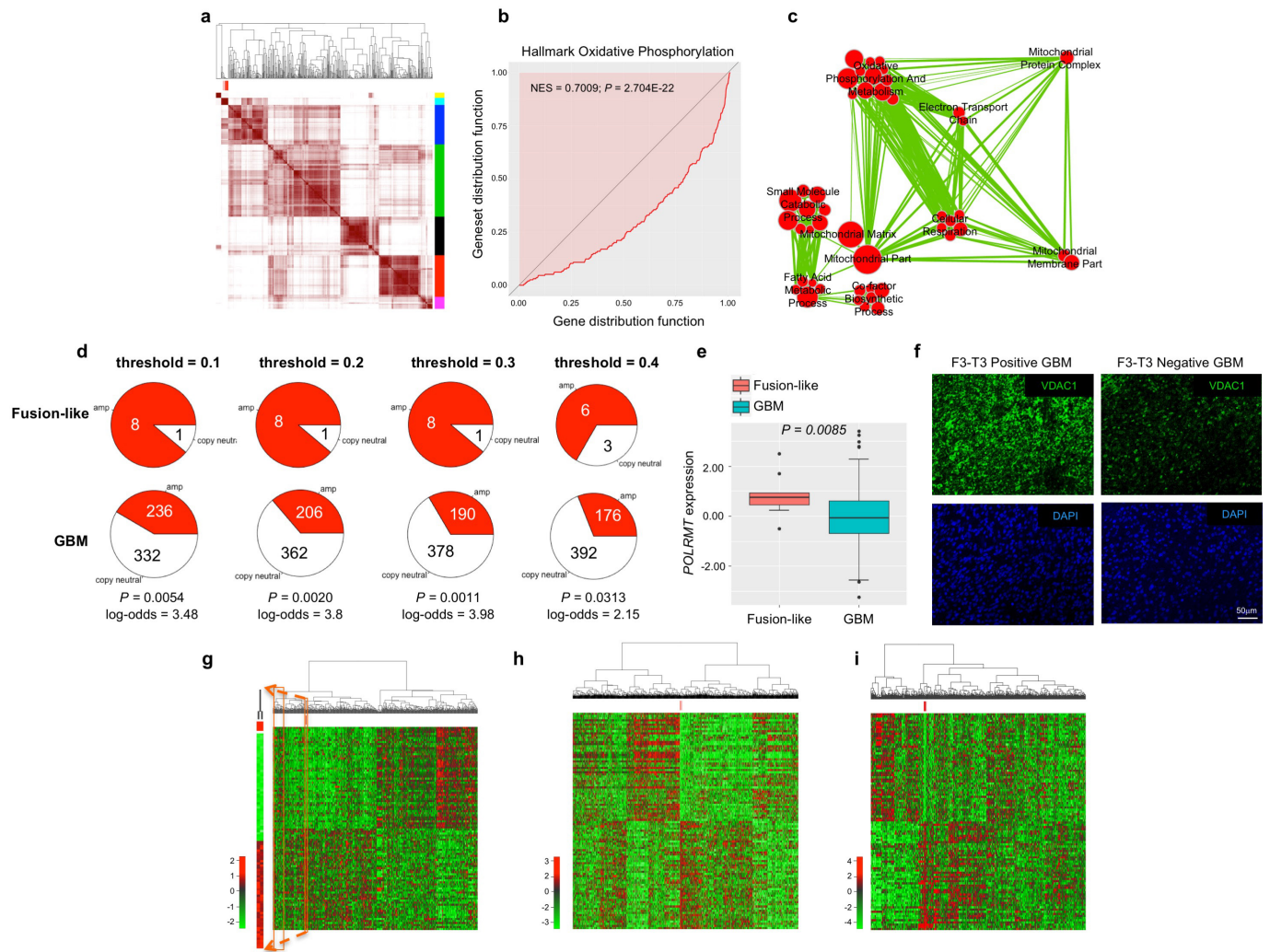
a, Amino acid sequence flanking Y122 of PIN4 (in red) is evolutionarily conserved. **b**, Immunoprecipitation-western blot analysis of human astrocytes expressing F3-T3 or F3-T3(K508M) with or without silencing of endogenous PIN4. β-Actin is shown as a loading control. WCL, whole-cell lysate. **c**, Immunoblot analysis of phosphotyrosine immunoprecipitates (left) or whole-cell lysates (right from U87 glioma cells expressing empty vector, FGFR3, F3-T3 or F3-T3(K508M) using the indicated antibodies. The asterisk indicates a non-specific band. **d**, Immunoblot analysis of phosphotyrosine immunoprecipitates (left) or whole-cell lysates (right) from human astrocytes expressing empty vector, F3-T3 or F3-T3(K508M) using the indicated antibodies. FAK is shown as a loading control. **e**, Immunoblot analysis of phosphotyrosine immunoprecipitates (left) or whole-cell lysates (right) from GSC1123 cells expressing endogenous F3-T3 shows decreased phosphorylation of F3-T3 substrates following treatment with AZD4547 for the indicated times. Paxillin is shown as a

loading control. **f**, Immunoblot analysis of canonical FGFR signalling proteins in GSC1123 cells treated with AZD4547 for the indicated time. β-Actin is shown as a loading control. **g**, Immunoblot analysis of phosphotyrosine immunoprecipitates from human astrocytes expressing F3-T3 or vector transduced with wild-type or the unphosphorylatable Y to A F3-T3 kinase substrate mutants. Paxillin is shown as a loading control. The asterisk indicates a non-specific band. **h**, Confocal images of immunofluorescence staining using the phospho-PIN4(Y122)-specific antibody (red) in human astrocytes transduced with vector or F3-T3 without or with silencing of endogenous PIN4. Nuclei were stained with DAPI (blue). **i**, Immunoblot analysis of phospho-PIN4(Y122), total PIN4 and FGFR3 in SF126 cells transduced with FGFR3, F3-T3, F3-T3(K508M) or the empty vector. β-Actin is shown as a loading control. Molecular weights are indicated in all panels. Experiments in **b-g, i** were repeated independently three times with similar results. Experiment in **h** was repeated independently four times with similar results.



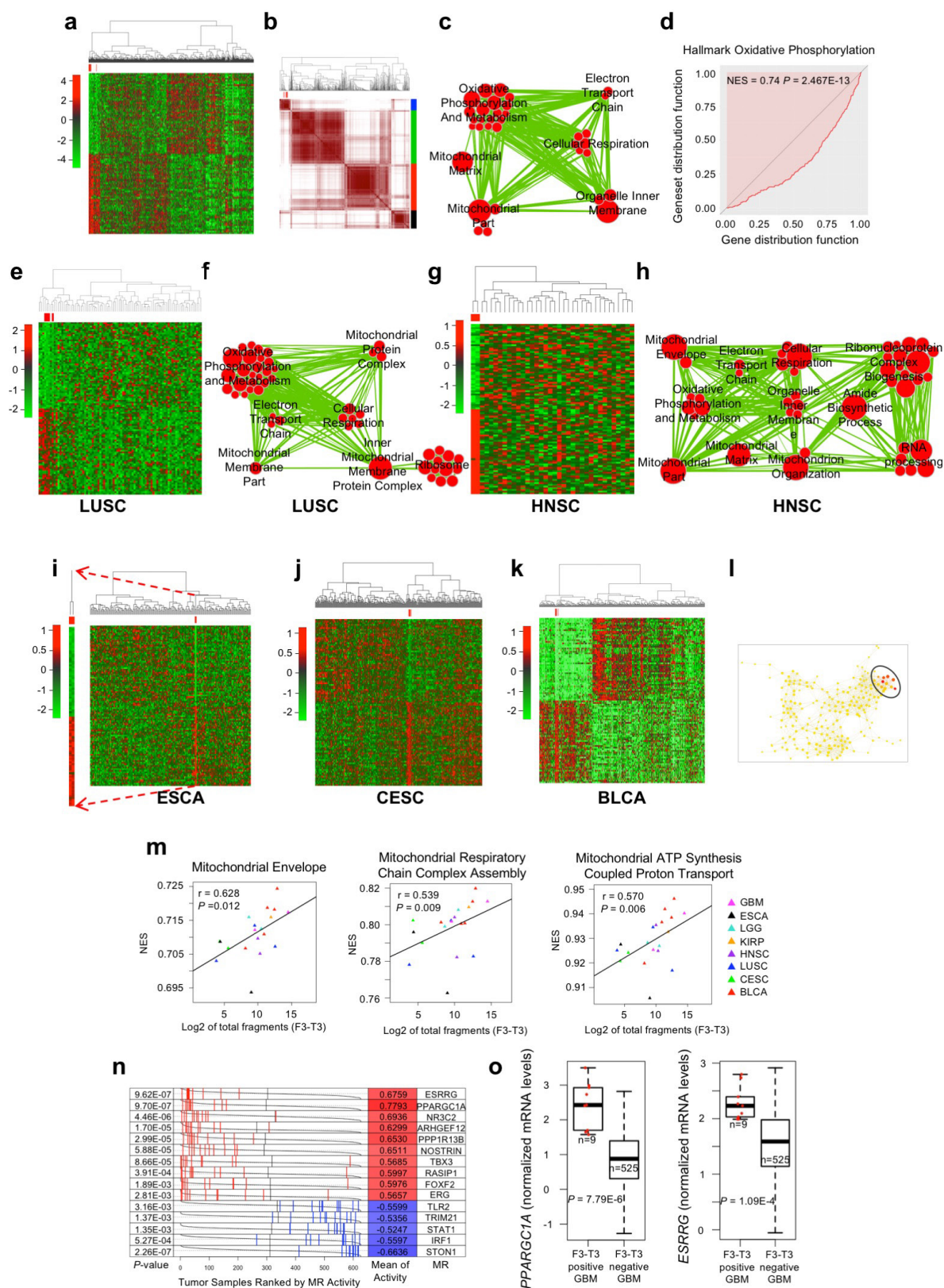
Extended Data Figure 4 | Functional analysis of tyrosine phosphorylation of F3-T3 kinase substrates. **a**, Western blot analysis of phosphotyrosine immunoprecipitation of F3-T3;shTrp53 and HRAS(12V);shTrp53 mGSCs using the PIN4 antibody. F3-T3 and HRAS(12V) expression are shown. α -Tubulin is shown as a loading control. **b**, Immunofluorescence images using the phospho-PIN4(Y122)-specific antibody (red, top) in tumours from F3-T3;shTrp53 and HRAS(12V)shTrp53 mGSCs. Nuclei were counterstained with DAPI (blue, bottom). Experiment was repeated independently twice with similar results. **c**, Left, representative images of phospho-PIN4(Y122) immunofluorescence in F3-T3-positive (top) and F3-T3-negative (bottom) GBM (green). Right, higher magnification images of phospho-PIN4(Y122)-DAPI co-staining depicting cytoplasmic localization of phospho-PIN4(Y122). Middle, DAPI staining of nuclei is shown as an indication of cellular density. **d**, Analysis of OCR in human astrocytes F3-T3 transduced with wild-type or the unphosphorylatable Y to A mutant of GOLGIN84, C1orf50 and DLG3. Human astrocytes expressing the empty vector are included as a control. Data are mean \pm s.d. ($n = 5$ technical replicates) of one representative experiment out of two independent experiments performed in triplicate with similar results. $P < 0.001$, rate 9–12 for vector versus each F3-T3 combination, two-tailed

t -test with unequal variance. **e**, Analysis of OCR of human astrocytes expressing F3-T3 transduced with PKM2(WT), PKM2(Y105A) or the empty vector. Human astrocytes expressing the empty vector are included as control. Data are mean \pm s.d. ($n = 3$ technical replicates) of one representative experiment out of three independent experiments; $P < 0.001$, rate 9–12 for vector versus each F3-T3 combination, two-tailed t -test with unequal variance. **f**, Immunoblot analysis of GOLGIN84, C1orf50 and DLG3 wild-type or Y to A mutants in human astrocytes expressing F3-T3 or vector. **g**, Immunoblot analysis of human astrocytes transduced with empty vector or F3-T3 expressing PKM2(WT) or PKM2(Y105A). **h**, Immunoblot analysis of human astrocytes transduced with F3-T3 or the empty vector for the expression of PIN4(WT) or PIN4(Y122F). **i**, Immunoblot analysis of PIN4 proteins in human astrocytes expressing F3-T3 following silencing of endogenous PIN4 and reconstitution with PIN4(WT), PIN4(Y122A) or PIN4(Y122F). In f–i, β -actin is shown as a loading control. Molecular weights are indicated on all immunoblots. **j**, Quantification of ATP levels in human astrocytes treated as in i. Data are mean \pm s.d. ($n = 4$ technical replicates) of one out of two independent experiments. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; two-tailed t -test with unequal variance. Experiments in **a**, **f**–**i** were repeated independently three times with similar results.



Extended Data Figure 5 | Transcriptomic analysis of F3-T3 fusion-positive and fusion-like GBM and validation of ee-MWW using different cancer-driving alterations. **a**, Consensus clustering on the Euclidean distance matrix based on the top and bottom 50 genes having the highest and lowest probability to be upregulated, respectively, in the nine F3-T3 fusion-positive GBM. The consensus matrix is obtained from 10,000 random samplings using 70% of the 544 samples. The nine F3-T3-positive GBM (in red) fall in one cluster (cyan). **b**, MWW enrichment plot of the 'hallmark oxidative phosphorylation' GO category in F3-T3-positive GBM. NES and P -values are indicated. **c**, Enrichment map network of statistically significant GO categories ($Q < 0.001$, NES > 0.6 , upper-tailed MWW-GST) in nine fusion-like GBM. Nodes represent GO terms and lines demonstrate their connectivity. Size of nodes is proportional to number of genes in the GO category and thickness of lines indicates the fraction of genes shared between the groups. **d**, Analysis of copy number amplification of the *POLRMT* gene comparing fusion-like GBM with all other GBM at different thresholds for amplification detection on log-R

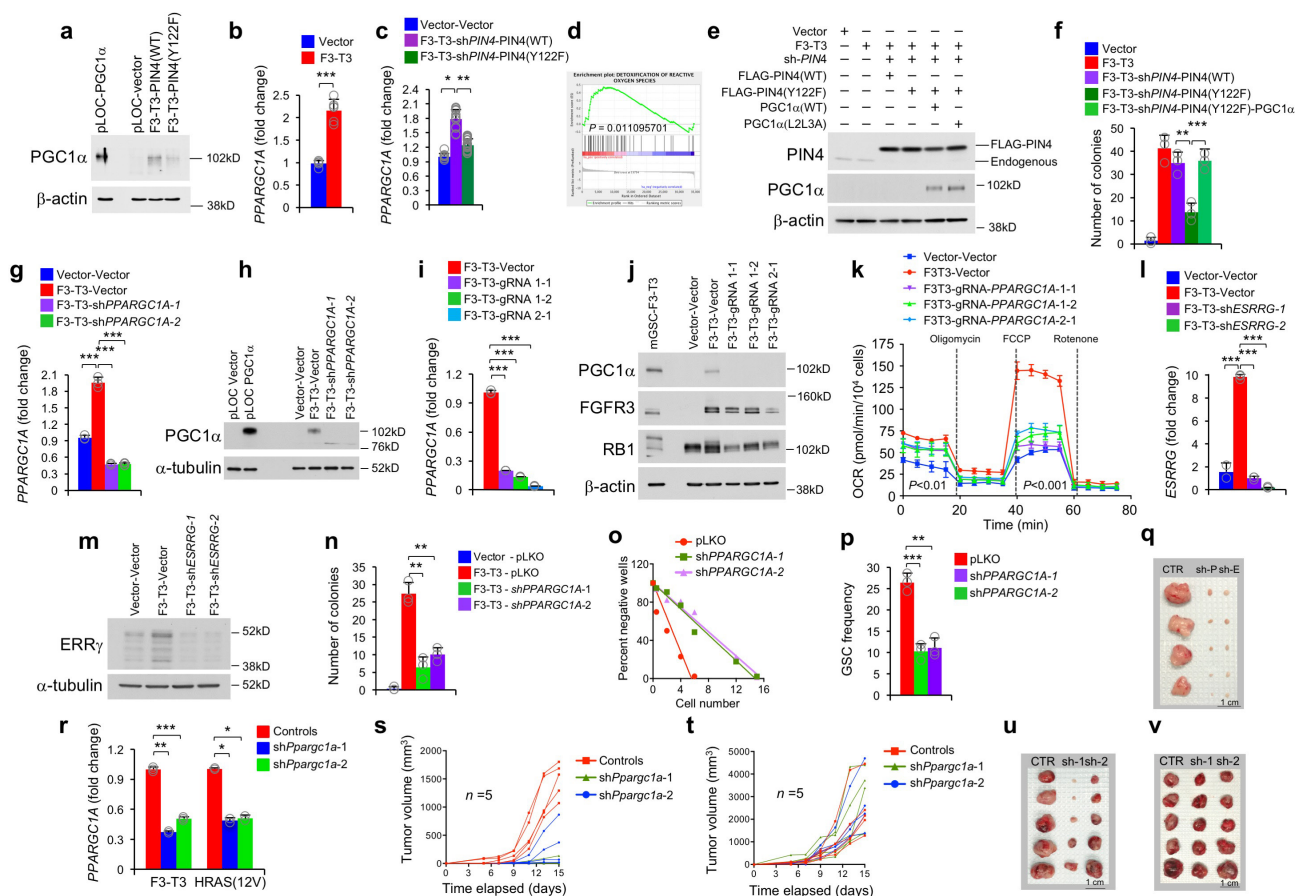
ratio from single-nucleotide polymorphism arrays. P value and log-odds at different thresholds are indicated (Fisher's exact test). **e**, *POLRMT* gene expression in fusion-like GBM ($n = 9$) and the remaining samples ($n = 535$). Box plot spans the first to third quartiles and whiskers show the $1.5 \times$ interquartile range. P value, two-sided MWW test. **f**, Representative images of VDAC1 immunofluorescence in F3-T3-positive (green, left) and F3-T3-negative (right) GBM. DAPI staining of nuclei is shown as an indication of cellular density (blue, bottom). **g**, Hierarchical clustering of two GBM samples with a KRAS mutation (red) out of 544 samples. Heat map of the two KRAS mutant samples is enlarged to the left. **h**, Hierarchical clustering of five KRAS-mutated samples (red) in the invasive breast carcinoma (BRCA) cohort ($n = 1,093$). **i**, Hierarchical clustering of six EGFR-SEPT14-positive GBM samples (red) out of 544 samples. Data in **g**–**i**, were obtained using the Euclidean distance and Ward linkage method and are based on the top and bottom 50 genes having the highest and lowest probability to be upregulated, respectively.



Extended Data Figure 6 | See next page for caption.

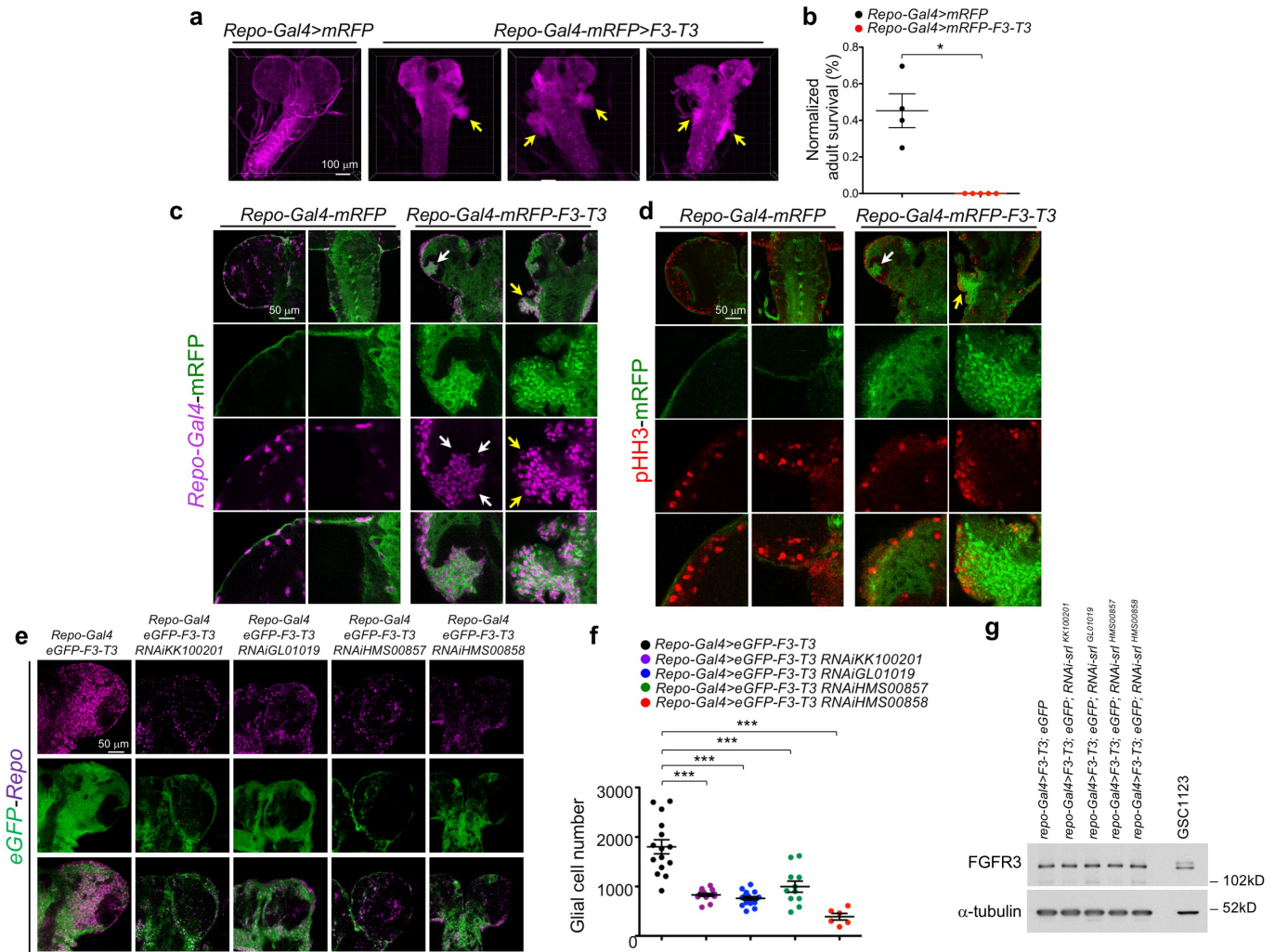
Extended Data Figure 6 | Pan-glioma and multi-cancer analysis of F3-T3 fusion-positive samples. **a, b**, Hierarchical (a) and consensus clustering (b) of 11 F3-T3-positive samples (red) out of 627 pan-glioma samples. The 11 F3-T3-positive samples (red) in **b** fall in one cluster (blue). **c**, Enrichment map network of statistically significant GO categories ($Q < 0.001$, NES > 0.6 ; upper-tailed MWW-GST) in the 11 F3-T3 fusion-positive pan-glioma samples. Nodes represent GO terms and lines demonstrate their connectivity. Size of nodes is proportional to number of genes in the GO category and thickness of lines indicates the fraction of genes shared between the groups. **d**, MWW enrichment plot of the 'hallmark oxidative phosphorylation' GO category in F3-T3-positive samples in the pan-glioma cohort. **e**, Hierarchical clustering of four F3-T3-positive (red) samples out of 86 lung squamous cell carcinoma (LUSC) samples. **f**, Enrichment map network of statistically significant GO categories ($Q < 0.001$, NES > 0.6 ; upper-tailed MWW-GST) in four F3-T3-positive LUSC. Nodes represent GO terms and lines demonstrate their connectivity. Size of nodes is proportional to number of genes in the GO category and thickness of lines indicates the fraction of genes shared between the groups. **g**, Hierarchical clustering of two F3-T3-positive, human papilloma virus (HPV)-positive head and neck squamous cell carcinoma (HNSC) samples (in red) out of 36 samples. **h**, Enrichment map network of statistically significant GO categories ($Q < 0.001$, NES > 0.6 ; upper-tailed MWW-GST) in two F3-T3-positive HNSC samples. Nodes represent GO terms and lines demonstrate their connectivity. Size of nodes is proportional to number of genes in the GO category and thickness of lines indicates the fraction of genes shared between the groups. **i**, Hierarchical clustering of two F3-T3-positive samples (red) out of 184 oesophageal carcinoma (ESCA) samples. Heat maps of the two

F3-T3-positive samples are enlarged to the left. **j**, Hierarchical clustering of four F3-T3-positive samples (red) out of 305 cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) samples. **k**, Hierarchical clustering of five F3-T3-positive samples (red) out of 408 urothelial bladder carcinoma (BLCA) samples. **l**, TDA network of pan-glioma samples ($n = 627$) reconstructed using variance normalized Euclidean distance and locally linear embedding as filter function. The nodes containing F3-T3-positive samples are highlighted in red. **m**, Correlation between the expression of F3-T3 (\log_2 of total fragment, x axis) and NES (y axis) of three top ranking mitochondrial functional categories in a multi-cancer cohort including F3-T3-positive samples ($n = 19$) from eight tumour types (r and P values are indicated, upper-tailed Spearman's rank correlation test). **n**, Analysis of the activity of master regulators in the pan-glioma cohort ($n = 627$ glioma). Grey curves represent the activity of each master regulator with tumour samples ranked according to master regulator activity. Red and blue lines indicate individual F3-T3-positive GBM samples displaying high and low master regulator activity, respectively. P values, two-sided MWW test, for differential activity (left) and mean of the activity (right) of the master regulator in F3-T3-positive versus F3-T3-negative samples are indicated. **o**, Gene expression analysis of *PPARGC1A* and *ESRRG* genes in F3-T3-positive and F3-T3-negative GBM; $n = 9$ F3-T3-positive tumours; $n = 525$ F3-T3-negative tumours. Box plot spans the first to third quartiles and whiskers show the $1.5 \times$ interquartile range. P value, two-sided MWW test. Data in **a**, **e**, **g**, **i-k**, were obtained using the Euclidean distance and Ward linkage method and are based on the top and bottom 50 genes having the highest and lowest probability to be upregulated, respectively.



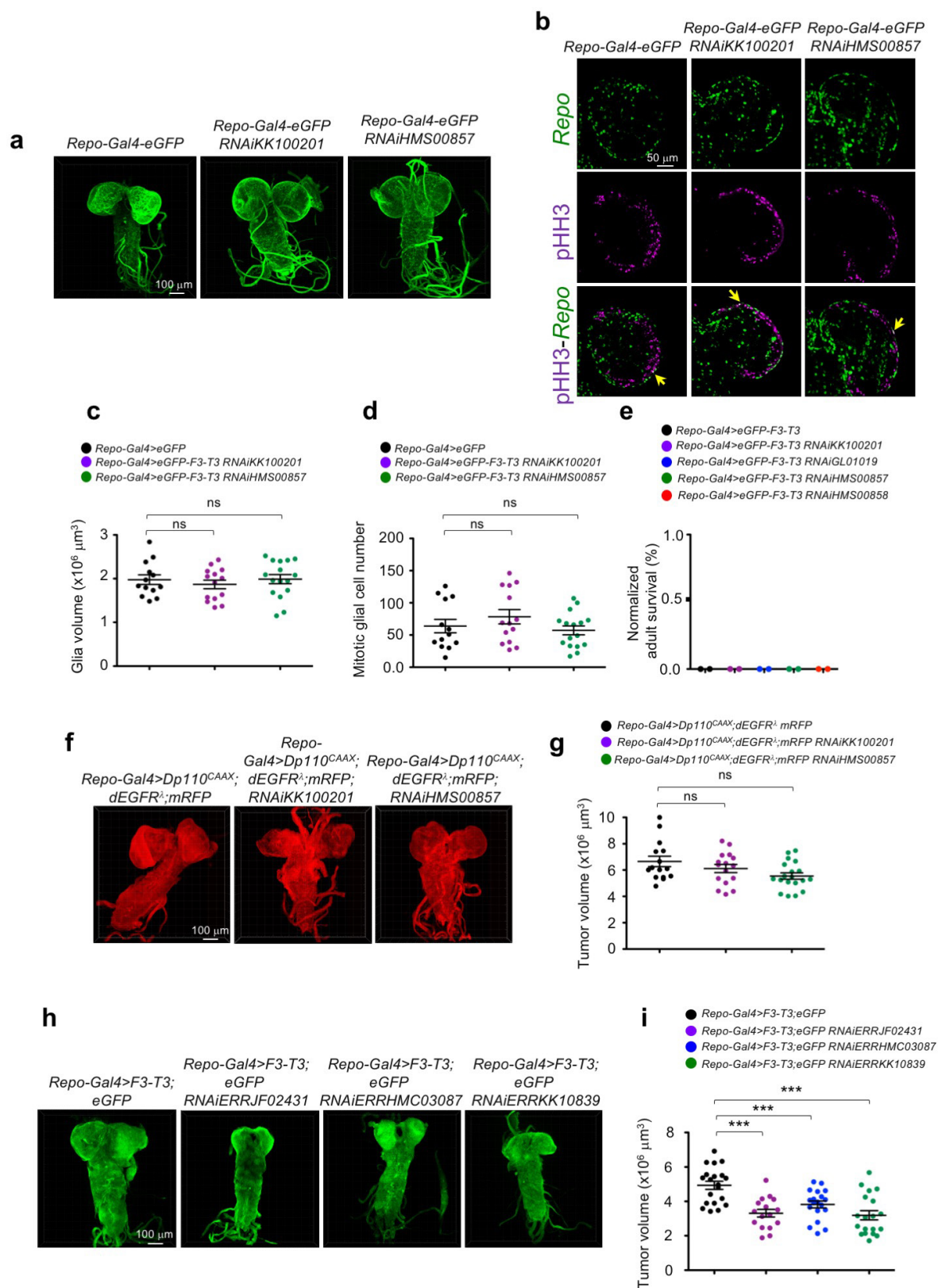
Extended Data Figure 7 | PGC1α and ERRγ are required for mitochondrial metabolism and tumorigenesis of cells transformed by F3-T3. **a**, Immunoblot of endogenous PGC1α in human astrocytes expressing F3-T3 following silencing of PIN4 and reconstitution with wild-type or PIN4(Y122F). Exogenous expression of PGC1α in human astrocytes is included as positive control. Experiment was independently repeated three times with similar results. **b**, RT-qPCR of *PPARGC1A* in human astrocytes expressing F3-T3 or vector. Data are mean \pm s.d. ($n = 6$ replicates) from two independent experiments each performed in triplicate. **c**, RT-qPCR of *PPARGC1A* in human astrocytes expressing F3-T3 treated as in **a**. Data are mean \pm s.d. ($n = 4$ biological replicates) from four independent experiments. **d**, GSEA shows upregulation of ROS detoxification genes in human astrocytes expressing F3-T3 ($n = 5$ biological replicates) compared with vector ($n = 3$ biological replicates). Nominal P value is indicated. **e**, Immunoblot of Flag-PIN4 (wild-type and Y122F) and PGC1α (wild-type and L2L3A) in human astrocytes expressing F3-T3 after silencing of PIN4. Experiment was repeated twice independently with similar results. **f**, Soft agar colony-forming assay of human astrocytes F3-T3 following silencing of PIN4 and reconstitution with wild-type or Y122F Flag-PIN4 in the presence or the absence of PGC1α. Data are mean \pm s.d. ($n = 3$ technical replicates) from one representative experiment out of two independent experiments. **g**, RT-qPCR of *PPARGC1A* in human astrocytes expressing vector or F3-T3 transduced with *PPARGC1A* shRNA1 or *PPARGC1A* shRNA2 lentivirus. Data are mean \pm s.d. ($n = 3$ technical replicates) from one representative experiment. **h**, Immunoblot analysis of PGC1α in HA-F3-T3 treated as in **g**. Experiment was repeated twice independently with similar results. Exogenous expression of PGC1α is included as positive control. Experiment was repeated twice independently with similar results. **i**, RT-qPCR of *PPARGC1A* in F3-T3 human astrocytes expressing two independent gRNAs against *PPARGC1A* (*PPARGC1A* gRNA1, two clones; *PPARGC1A* gRNA2, 1 clone) or the empty vector. Data are mean \pm s.d. ($n = 3$ technical replicates) from one representative experiment. **j**, Western blot of cells treated as in **(i)** using the indicated antibodies. Experiment was repeated twice independently with similar results. **k**, OCR of human astrocytes expressing vector or F3-T3 transduced with *PPARGC1A* gRNA1 or gRNA2. Data are mean \pm s.d. ($n = 5$ technical replicates) from one representative experiment out of two independent experiments. $P < 0.001$

for rate 1–4 and 9–12; two-tailed t -test with unequal variance. **l**, RT-qPCR of *ESRRG* in human astrocytes expressing vector or F3-T3 infected with *ESRRG* shRNA1 or *ESRRG* shRNA2 lentiviruses. Data are mean \pm s.d. ($n = 3$ technical replicates) from one representative experiment. **m**, Immunoblot analysis of ERRγ in human astrocytes expressing F3-T3 treated as in **l**. Experiment was repeated twice independently with similar results. **n**, Soft agar colony-forming assay of human astrocytes treated as in Fig. 3g. Data are mean \pm s.d. ($n = 3$ technical replicates) of one representative experiment out of two independent experiments performed in triplicate. **o**, GSC1123 cells were transduced with *PPARGC1A* shRNA lentiviruses or the empty vector. Cells were analysed by *in vitro* LDA. Representative regression plot used to calculate the frequency of gliomaspheres in 96-well cultures from three independent infections. **p**, Bar graph shows the frequency of gliomaspheres from three independent infections analysed by LDA as shown in **o**. Data are mean \pm s.d. ($n = 3$ biological replicates). **q**, The photograph shows tumours generated by human astrocytes F3-T3 transduced with *PPARGC1A* shRNA1, *ESRRG* shRNA1 or vector lentivirus in Fig. 3i at the time of mouse euthanasia. sh-P, *PPARGC1A* shRNA1; sh-E, *ESRRG* shRNA1. **r**, RT-qPCR of *Ppargc1a* in F3-T3-shTrp53 and HRAS(12V) shTrp53 mGSCs transduced with *Ppargc1a* shRNA1 or *Ppargc1a* shRNA2 lentivirus. Data are mean \pm s.d. ($n = 3$ technical replicates) of one representative experiment. **s**, Tumour volume of F3-T3-shTrp53 mGSCs expressing a pLKO-vector ($n = 5$), *Ppargc1a* shRNA1 ($n = 5$) or *Ppargc1a* shRNA2 ($n = 5$). Data are the tumour growth curve of individual mice. **t**, Tumour volume of mice injected subcutaneously with HRAS(12V);shTrp53 mGSCs expressing pLKO-vector ($n = 5$) or *Ppargc1a* shRNA1 ($n = 5$) or *Ppargc1a* shRNA2 ($n = 5$). Data are tumour growth curve of individual mice; NS, not significant, two-tailed t -test with unequal variance (time points 1–7). **u**, Photograph shows tumours generated from F3-T3;shTrp53 mGSCs transduced with *Ppargc1a* shRNA1 or *Ppargc1a* shRNA2 or vector lentivirus in **s** at the time of mouse euthanasia. **v**, Photograph shows tumours generated by HRAS(12V) shTrp53 mGSCs transduced with *Ppargc1a* shRNA1 or *Ppargc1a* shRNA2 or vector lentivirus in **t** at the time of mouse euthanasia. Molecular weights are indicated and β -actin or α -tubulin is shown as a loading control in all immunoblots. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; two-tailed t -test with unequal variance.



Extended Data Figure 8 | *Drosophila* PGC1 α -homologue *spargel* (*srl*) mediates F3-T3-induced tumour growth. **a**, Optical projections of whole brain-ventral nerve cord complexes from *Drosophila* larvae. Expression of the F3-T3 fusion oncogene using the *repo-Gal4* (*repo-Gal4>F3-T3*) pan-glial driver induced pathological changes in brain and ventral nerve cord with ectopic tissue protrusions (yellow arrows) due to excessive glial cell proliferation and accumulation. **b**, Survival of larvae bearing F3-T3-driven glial tumours. Larvae bearing F3-T3-driven glial tumours die before developing into adulthood (biologically independent samples: $n = 87$, *Repo-Gal4>mRFP*; $n = 77$, *Repo-Gal4>mRFP-F3-T3*). Data are shown as mean \pm s.e.m. * $P < 0.05$; two-tailed t -test with unequal variance. Individual dots represent the fraction of surviving animals. **c**, Glial expression of F3-T3 resulted in increased total glial cell number (*Repo*⁺*mRFP*⁺ cells) compared to controls. Note the excessive accumulation of glial cells in the brain lobe (white arrows) and ventral nerve cord (yellow arrows). **d**, Glial expression of F3-T3 increases glial

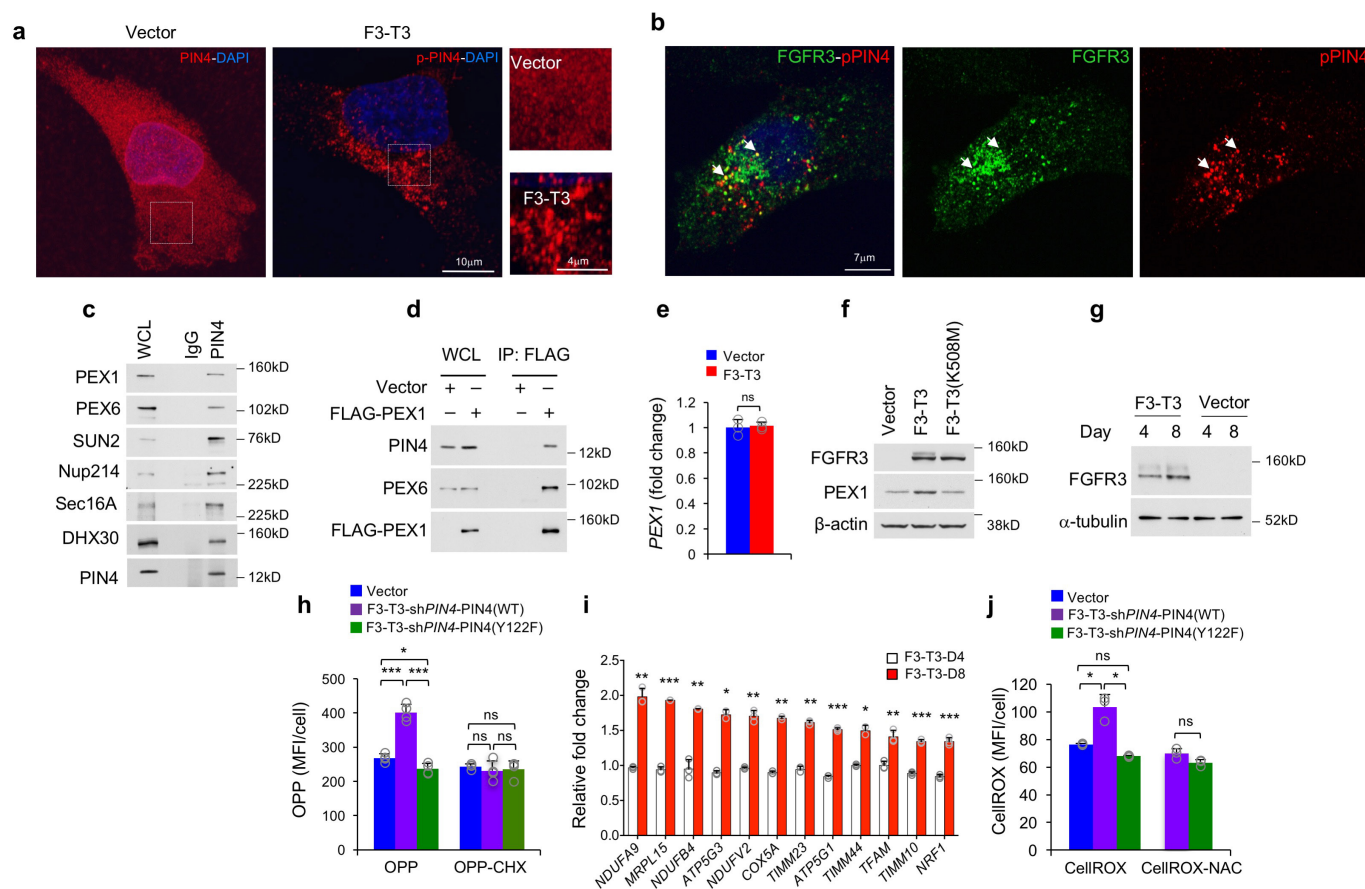
cell proliferation (*mRFP*⁺ phosphorylated histone H3⁺ (phospho-HH3⁺) cells) compared to control. Note the excessive accumulation of glial cells in the brain lobe (white arrows) and ventral nerve cord (yellow arrows). **e**, Glia-specific *srl* knockdown in F3-T3-induced glial tumours resulted in decreased total glial number (*Repo*⁺*eGFP*⁺ cells) compared to controls. **f**, Quantification of glia number in control and *srl*-deficient tumours. $n = 15$ for *repo-Gal4>F3-T3*; $n = 15$ for *repo-Gal4>F3-T3* *RNAi-KK100201*; $n = 16$ for *repo-Gal4>F3-T3*; *RNAi-GL01019*; $n = 11$, for *repo-Gal4>F3-T3*; *RNAi-HMS00857*; $n = 6$ for *repo-Gal4>F3-T3*; *RNAi-HMS00858*. Data are shown as mean \pm s.e.m. *** $P < 0.001$; two-tailed t -test with unequal variance. **g**, Western blot analysis of the F3-T3 protein in *repo-Gal4>F3-T3* and *repo-Gal4>F3-T3*; *RNAi-srl* *Drosophila* brains. The expression of F3-T3 in human GSC1123 cells is shown as a positive control for F3-T3 and α -tubulin is shown as a loading control. Experiments in **c-e**, **g** were performed twice.



Extended Data Figure 9 | See next page for caption.

Extended Data Figure 9 | Glia-specific knockdown of *srl* has little to no effect on EGFR–PI3K-induced tumour growth but glia-specific knockdown of *ERR* inhibits F3–T3-induced tumour growth. **a**, Optical projections of whole brain–ventral nerve cord complexes from larvae with control and *srl*-deficient glia. **b**, Glia-specific *srl* knockdown in larval brains did not significantly affect the overall glial population (Repo^+ cells) nor the mitotic index of glial cells (Repo^+ phospho-HH3⁺ cells, yellow arrows). **c**, Quantification of glia volume in larval brains with control and *srl*-deficient glia; $n = 13$ for *repo-Gal4>eGFP*; $n = 14$ for *repo-Gal4>eGFP;RNAi-KK100201*; and $n = 16$, for *repo-Gal4>eGFP;RNAi-HMS00857*. Data are mean \pm s.e.m. NS, not significant; two-tailed *t*-test with unequal variance. **d**, Quantification of proliferating glia number (Repo^+ ; phospho-HH3⁺ cells) in larval brains with control and *srl*-deficient glia; $n = 13$ for *repo-Gal4>eGFP*; $n = 14$ for *repo-Gal4>eGFP;RNAi-KK100201*; and $n = 16$ for *repo-Gal4>eGFP;RNAi-HMS00857*. Data are mean \pm s.e.m. NS, not significant; two-tailed *t*-test with unequal variance. **e**, Adult lethality in

repo-Gal4>F3–T3 and *repo-Gal4>F3–T3;RNAi-srl* larvae ($n > 100$). **f**, Optical projections of control and *srl*-deficient brain tumours from *repo-Gal4>Dp110^{CAAX};dEGFR^Δ;mRFP* larvae. **g**, Quantification of tumour volume in control and *srl*-deficient tumours; $n = 15$ for *repo-Gal4>Dp110^{CAAX};dEGFR^Δ;mRFP*; $n = 16$ for *repo-Gal4>Dp110^{CAAX};dEGFR^Δ;mRFP;RNAi-KK100201*; $n = 19$ for *repo-Gal4>Dp110^{CAAX};dEGFR^Δ;mRFP;RNAi-HMS00857*. Data are mean \pm s.e.m. NS, not significant; two-tailed *t*-test with unequal variance. **h**, Optical projections of brain tumours from *Drosophila* larvae *repo-Gal4>F3–T3* and *repo-Gal4>F3–T3;RNAi-ERR*. RNAi-mediated knockdown of *ERR* reduces the volume of F3–T3-induced glial tumours. **i**, Quantification of tumour volume in the control and *ERR*-deficient tumours; $n = 20$ for *repo-Gal4>F3–T3*; $n = 16$ for *repo-Gal4>F3–T3;RNAi-JF02431*; $n = 19$, for *repo-Gal4>F3–T3;RNAi-HMC03087*; $n = 19$, for *repo-Gal4>F3–T3;RNAi-KK10839*. *** $P < 0.001$; two-tailed *t*-test with unequal variance. In all experiments n are biologically independent animals. Experiments in **a**, **b**, **f**, **h**, were repeated twice with similar results.



Extended Data Figure 10 | Acute expression of F3-T3 fusion induces peroxisome biogenesis through phosphorylation of PIN4(Y122). **a**, Representative confocal images (maximum intensity) of immunofluorescence staining for total PIN4 (PIN4, red, left) and phospho-PIN4(Y122) (p-PIN4, red, middle panel) in human astrocytes expressing the empty vector and F3-T3. Right, higher magnification of dotted boxes. Nuclei were counterstained with DAPI (blue). Experiment was repeated independently twice with similar results. **b**, Maximum intensity of confocal images of double immunofluorescence staining for FGFR3 (green, middle) and phospho-PIN4(Y122) (red, right) in human astrocytes expressing F3-T3. Arrows indicate protein co-localization. Experiment was repeated independently twice with similar results. **c**, Co-immunoprecipitation from H1299 cells using the PIN4 antibody. Endogenous PIN4 immunocomplexes and input (WCL) were analysed by western blot using the indicated antibodies. Input is 10% for PEX1, PEX6, SUN2 and NUP214; 5% for SEC16A and DHX30; 2% for PIN4. **d**, Western blot analysis of co-immunoprecipitation of exogenous Flag-PEX1 in human astrocytes expressing F3-T3. WCL: 1% for PIN4 and 10% for PEX1 and PEX6. Experiment was repeated independently four times with similar results. **e**, RT-qPCR of *PEX1* in human astrocytes expressing F3-T3 or vector. Data are mean \pm s.d. ($n = 3$ technical replicates) of one representative experiment out of three independent experiments performed in triplicate. **f**, Western blot analysis of PEX1 expression in human astrocytes transduced with F3-T3, F3-T3(K508M) or the empty vector. β -Actin is shown as a loading control. Experiment was repeated

independently three times with similar results. **g**, Time-course analysis of F3-T3 expression in human astrocytes by western blot. α -Tubulin is shown as a loading control. Experiment was repeated independently twice with similar results. **h**, Quantification of protein biosynthesis by OPP incorporation measured by high-content fluorescent microscopy in human astrocytes reconstituted with PIN4(WT) or PIN4(Y122F) after silencing of the endogenous *PIN4* and acutely transduced with F3-T3 or vector. Representative bar plots ($n = 4$ technical replicates) from one out of three independent experiments. * $P < 0.05$, *** $P < 0.001$; two-tailed t -test with unequal variance. CHX-treated cultures were used as negative controls. **i**, Time-course expression analysis by RT-qPCR of the indicated mitochondrial genes in human astrocytes expressing F3-T3 or empty vector. Data are mean \pm s.d. ($n = 3$ technical replicates) of one representative experiment out of two independent experiments performed in triplicate. Values were normalized to vector (dotted line). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; two-tailed t -test with unequal variance. **j**, Quantification of cellular ROS (measured by high-content microscopy) in human astrocytes reconstituted with PIN4(WT) or PIN4(Y122F) after silencing of the endogenous *PIN4* and acutely transduced with F3-T3 or vector. Representative bar plots from one out of three independent experiments. Data are mean \pm s.d. ($n = 3$ technical replicates). * $P < 0.05$; two-tailed t -test with unequal variance. N-acetyl-L-cysteine-treated cultures were used as negative controls. Molecular weights are indicated in all immunoblots.

Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators

Nian Liu^{1*}, Cameron H. Lee^{2*}, Tomek Swigut¹, Edward Grow^{2,7}, Bo Gu¹, Michael Bassik^{2,3} & Joanna Wysocka^{1,4,5,6}

Transposable elements (TEs) are now recognized not only as parasitic DNA, whose spread in the genome must be controlled by the host, but also as major players in genome evolution and regulation^{1–6}. Long INterspersed Element-1 (LINE-1 or L1), the only currently autonomous mobile transposon in humans, occupies 17% of the genome and continues to generate inter- and intra-individual genetic variation, in some cases resulting in disease^{1–7}. Nonetheless, how L1 activity is controlled and what function L1s play in host gene regulation remain incompletely understood. Here, we use CRISPR/Cas9 screening strategies in two distinct human cell lines to provide the first genome-wide survey of genes involved in L1 retrotransposition control. We identified functionally diverse genes that either promote or restrict L1 retrotransposition. These genes, often associated with human diseases, control the L1 lifecycle at transcriptional or post-transcriptional levels and in a manner that can depend on the endogenous L1 sequence, underscoring the complexity of L1 regulation. We further investigated L1 restriction by MORC2 and human silencing hub (HUSH) complex subunits MPP8 and TASOR⁸. HUSH/MORC2 selectively bind evolutionarily young, full-length L1s located within transcriptionally permissive euchromatic environment, and promote H3K9me3 deposition for transcriptional silencing. Interestingly, these silencing events often occur within introns of transcriptionally active genes and lead to down-regulation of host gene expression in a HUSH/MORC2-dependent manner. Together, we provide a rich resource for studies of L1 retrotransposition, elucidate a novel L1 restriction pathway, and illustrate how epigenetic silencing of TEs rewires host gene expression programs.

Most of our knowledge about L1 retrotransposition control comes from studies examining individual candidate genes^{2–6}. To systematically identify genes regulating L1 retrotransposition, we performed a genome-wide CRISPR/Cas9 screen in human chronic myeloid leukemia K562 cells using an L1-G418^R retrotransposition reporter⁹ (Fig. 1a,b). Importantly, the L1-G418^R reporter was modified to be driven by a doxycycline (dox)-responsive promoter, as opposed to the native L1 5'UTR, to avoid leaky retrotransposition ahead of the functional screen (Extended Data Fig. 1a–c). The cells become G418^R antibiotic resistant only when the L1-G418^R reporter undergoes a successful retrotransposition event following dox-induction (Fig. 1b). For the screen, we transduced clonal L1-G418^R cells with a lentiviral genome-wide sgRNA library such that each cell expressed a single sgRNA¹⁰. We then dox-induced the cells to turn on the L1-G418^R reporter for retrotransposition, and split the cells into G418-selected conditions and unselected conditions, which served to eliminate cell growth bias in the screen analysis. The frequencies of sgRNAs in the two populations were measured by deep sequencing (Fig. 1a) and analyzed using Cas9 high-Throughput maximum Likelihood Estimator

(CasTLE)¹¹. Consequently, cells transduced with sgRNAs targeting L1 suppressors would have more retrotransposition events than negative control cells and would be enriched through the G418 selection; conversely, cells transduced with sgRNAs targeting L1 activators would be depleted.

Using the above strategy, we identified 25 putative L1 regulators at a 10% FDR cutoff, and 150 genes at a 30% FDR cutoff (Fig. 1c and Extended Data Fig. 1d; see Table S1 for full list). Despite low statistical confidence, many of the 30% FDR cutoff genes overlapped previously characterized L1 regulators (e.g. ALKBH1, SETDB1) and genes functioning in complexes with our top 10% FDR hits (e.g. Fanconi Anemia pathway, HUSH complex), suggesting that they likely encompassed biologically relevant hits. To increase statistical power in distinguishing bona fide L1 regulators among these, we performed a high-coverage secondary screen targeting the 30% FDR hits (150 genes) and an additional 100 genes that were either functionally related to our top hits or which were otherwise previously known to regulate L1 but fell outside of the 30% FDR cutoff threshold (See Table S2 for full list). This secondary screen validated 90 genes out of the top 150 genome-wide screen hits, a fraction close to expected with the 30% FDR cutoff (Fig. 1d and Extended Data Fig. 2a–c).

Altogether, our two-tier screening approach identified 142 human genes that either activate or repress L1 retrotransposition in K562 cells, encompassing over 20 previously known L1 regulators (Extended Data Fig. 2d). Novel candidates are involved in functionally diverse pathways, such as chromatin/transcriptional regulation, DNA damage/repair, and RNA processing (Extended Data Fig. 2e,f). While many DNA damage/repair factors, particularly the Fanconi Anemia (FA) factors, suppress L1 activity, genes implicated in the Non-Homologous End Joining (NHEJ) repair pathway promote L1 retrotransposition (Extended Data Fig. 2f). In agreement, mutations in some of the identified NHEJ factors were previously found to result in decreased retrotransposition frequencies¹². Intriguingly, many hits uncovered by our screen (e.g. FA factors, MORC2 and SETX) are associated with human disorders^{13–17}.

To extend our survey of L1 regulators to another cell type, we performed both a genome-wide and a secondary screen in HeLa cells (Extended Data Fig. 1b, 1e) with the same sgRNA libraries used in the K562 screens. Importantly, top hits identified in the K562 genome-wide screen were recapitulated in the HeLa screen (e.g. MORC2, TASOR, SETX, MOV10) (Extended Data Fig. 3a). Furthermore, secondary screens in both K562 and HeLa cells showed concordant effects for groups of genes, for example, the suppressive effects of the FA complex genes, and activating effects of the NHEJ pathway genes (Extended Data Fig. 3b–e). Interestingly, however, a subset of genes showed cell-line selective effects (Extended Data Fig. 3c). At the same time, some of the previously known L1 regulators did not come up as hits in our

¹Department of Chemical and Systems Biology, Stanford School of Medicine, Stanford University, Stanford, CA 94305, USA. ²Department of Genetics, Stanford School of Medicine, Stanford University, Stanford, CA 94305, USA. ³Stanford University Chemistry, Engineering, and Medicine for Human Health (ChEM-H), Stanford School of Medicine, Stanford University, Stanford, CA 94305, USA. ⁴Institute of Stem Cell Biology and Regenerative Medicine, Stanford School of Medicine, Stanford University, Stanford, CA 94305, USA. ⁵Department of Developmental Biology, Stanford School of Medicine, Stanford University, Stanford, CA 94305, USA. ⁶Howard Hughes Medical Institute, Stanford School of Medicine, Stanford University, Stanford, CA 94305, USA.

⁷Present address: Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 841125550, USA.

*Co-first author

screen. Several factors could have limited our ability to identify all genes controlling L1 retrotransposition to saturation, such as: (i) a subset of regulators may function in a cell-type specific manner not captured by either K562 or HeLa screens, (ii) essential genes with strong negative effects on cell growth may have dropped out, (iii) regulators that strictly require native L1 UTR sequences may have been missed due to our reporter design. Nonetheless, our combined screens identify many novel candidates for L1 retrotransposition control in human cells and provide a rich resource for mechanistic studies of TEs.

Select screen hits were further validated in K562 cells using a well-characterized L1-GFP reporter¹⁸ (Extended Data Fig. 1a), confirming 13 suppressors and 1 activator (SLTM) out of 16 examined genes (Fig. 1e). Interestingly, chromatin regulators (TASOR, MORC2, MPP8, SAFB and SETDB1) suppress the retrotransposition of L1-GFP reporter, but not that of a previously described codon-optimized L1-GFP reporter (hereinafter referred to as (opt)-L1-GFP)^{19,20}, indicating that these factors regulate L1 retrotransposition in a manner dependent upon the native L1 ORF nucleotide sequence (Extended Data Fig. 3f,g). An additional secondary screen against the codon-optimized (opt)-L1-G418^R reporter in K562 cells confirmed the sequence-dependent feature of these L1 regulators, and systematically partitioned our top screen hits into native L1 sequence-dependent and -independent candidates (Extended Data Fig. 3h, see Table S2 for full list).

We next examined whether the identified regulators influence the expression of endogenous L1Hs, the youngest and only retrotransposition-competent L1 subfamily in humans. CRISPR-deletion of some genes (TASOR, MPP8, SAFB and MORC2) significantly increased expression of endogenous L1Hs, whereas deletion of other genes, such as SETX, RAD51 or FA complex components, had little effect (Fig. 1f). Since all interrogated genes restrict L1-GFP retrotransposition into the genome (Fig. 1e and Extended Data Fig. 4a), our results suggest that identified suppressors can function at either transcriptional or posttranscriptional level.

We further investigated three candidate transcriptional regulators of L1: MORC2, TASOR and MPP8. TASOR and MPP8 (along with PPHLN1), comprise the HUSH complex and recruit the H3K9me3 methyltransferase SETDB1 to repress genes⁸. Notably, PPHLN1 and SETDB1 also came up as L1 suppressors in our screen (Fig. 1d and Extended Data Fig. 3b). MORC2, which has recently been shown to biochemically and functionally interact with HUSH²¹, is a member of the microorchidia (MORC) protein family that has been implicated in transposon silencing in plants and mice^{22,23}. While MORC2/HUSH have been previously implicated in heterochromatin formation, most heterochromatin factors had no impact on L1 retrotransposition, suggesting a selective effect (Fig. 2a and Extended Data Fig. 4b).

Several independent experiments in clonal knockout (KO) K562 lines confirmed that HUSH and MORC2 suppress the retrotransposition of the L1-GFP reporter by silencing its transcription (Fig. 2b,c and Extended Data Fig. 4c-f). Additionally, HUSH/MORC2 repressed endogenous (non-reporter) L1Hs RNA and protein expression in both K562 and human embryonic stem cells²⁴ (hESC, H9) (Fig. 2d and Extended Data Fig. 4g-k). PolyA-selected RNA sequencing (RNA-seq) experiments revealed up-regulated expression of evolutionarily younger L1PA families (including L1Hs) upon HUSH or MORC2 KO in K562 cells (Fig. 2e). Taken together, these data demonstrate that HUSH/MORC2 silence both the reporter transgene as well as endogenous evolutionarily young L1s.

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) from K562 cells and hESCs demonstrated that MORC2, MPP8 and TASOR co-bind genomic regions characterized by specific L1 instances. Elements from the primate-specific L1P family showed higher enrichment than the older L1M family elements (Fig. 3a,b and Extended Data Fig. 5a,b, 7a,b), consistent with the preferential derepression of the former upon HUSH or MORC2 KO (Fig. 2e). Moreover, this enrichment was specific to L1s, as other major repeat classes were

not enriched (Fig. 3b and Extended Data Fig. 7b), although all three proteins also targeted expressed KRAB-ZNF genes (Extended Data Fig. 5c,d). HUSH KO in K562 cells almost completely abrogated MORC2 binding at L1s (consistent with recently published observations that HUSH recruits MORC2 for transcriptional repression²¹), whereas MORC2 deletion led to a modest, but appreciable decrease of HUSH subunit binding (Extended Data Fig. 6). In mouse ESCs, MPP8 bound retrotransposition-competent L1Md-A and L1Md-T, as well as IAP elements, a class of murine endogenous retroviruses that remain currently mobile in the mouse genome (Extended Data Fig. 7c,d), suggesting that regulators uncovered by our study in human cells may in other species target additional active transposons beyond L1s.

Interestingly, even within younger human L1Ps only a subset is bound by HUSH/MORC2 in either K562 cells or hESCs, and we sought to identify genomic or epigenomic features that could explain this selectivity. We found that HUSH/MORC2 selectively target young full-length L1s, particularly the L1PA1-5 in human cells (Fig. 3c,d) and L1Md-A/T in mice (Extended Data Fig. 7e). Both MPP8 and MORC2 bind broadly across the L1: while MORC2 binding is skewed towards the 5' end, MPP8 shows higher enrichments within the body and at 3' end of L1PAs, including the L1Hs (L1PA1) elements (Extended Data Fig. 7f,g).

Nonetheless, preference for the full-length, evolutionarily younger L1PAs can only partially explain observed HUSH/MORC2 selectivity, as only a subset of such elements is targeted by the complex (Fig. 3d). We found that the additional layer of selectivity can be explained by the state of surrounding chromatin, with HUSH/MORC2-occupied L1s preferentially immersed within the transcriptionally permissive euchromatic environment marked by modifications such as H3K4me3 and H3K27ac (Fig. 3e). In agreement, HUSH/MORC2-bound L1s are enriched within introns of actively transcribed genes (Extended Data Fig. 8a,b). Furthermore, although most HUSH/MORC2-bound L1s are concordant between K562 and hESCs, those that are bound in a cell type-specific manner tend to be associated with genes that are differentially active between the two cell types (Extended Data Fig. 8c). To understand the role of transcription in HUSH/MORC2 targeting of L1s, we investigated MORC2 and MPP8 occupancy at the inducible L1 transgene. We observed increased binding of these factors upon transcriptional induction (Extended Data Fig. 8d), suggesting that transcription through L1 sequences facilitates HUSH/MORC2 binding. Taken together, HUSH/MORC2 selectively target young, full-length L1s located within transcriptionally permissive euchromatic regions, which are precisely the elements that pose the highest threat to genome integrity, as a subset of them remains mobile and transcription is the first step of L1 mobilization.

Despite their immersion within the euchromatic environment, HUSH/MORC2-bound L1s themselves are heavily decorated with the transcriptionally repressive H3K9me3 (Fig. 3e), consistent with the role of HUSH in facilitating H3K9me3 deposition at target sites⁸. HUSH/MORC2 KO decreased H3K9me3 level preferentially at L1 versus non-L1 HUSH/MORC2 genomic targets, and at bound versus unbound L1s (Fig. 4a and Extended Data Fig. 9a,b). Since HUSH/MORC2-bound L1s are significantly enriched within introns of transcriptionally active genes (Extended Data Fig. 8a-c), we examined whether HUSH/MORC2 recruitment and its associated H3K9me3 deposition can influence chromatin modification and expression of the host genes. Despite the transcriptionally active status (Extended Data Fig. 8a,b), promoters and especially bodies of genes harboring MORC2/HUSH-bound L1s show appreciable levels of H3K9me3. This enrichment is substantially diminished in the KO lines (Extended Data Fig. 9c) with the concomitant upregulation of genes harboring MORC2/HUSH-bound L1s, but not those with unbound intronic L1s (Fig. 4b). Thus, HUSH/MORC2 binding at intronic L1s leads to a modest, but significant down-regulation of the active genes that harbor them (Fig. 4c and Extended Data Fig. 9d-g, 10a).

Inserting L1 sequences on a transcript leads to decrease in RNA expression via inadequate transcript elongation,²⁵ and this effect has been attributed to the A/T enrichment of L1s. However, our results argue that transcriptional attenuation of host gene expression could be a consequence of epigenetic silencing by HUSH/MORC2 (Fig. 4b,c and Extended Data Fig. 9d–g, 10a), and this possibility is consistent with the described role of genic H3K9me3 in decreasing Pol II elongation rate, leading to its accumulation over the H3K9me3 region²⁶. If such mechanism is at play, then HUSH KO should decrease accumulation of the elongating Pol II over L1 bodies, and this is indeed what we observe in Pol II ChIP-seq experiments (though interestingly, at 5' UTRs of L1s, Pol II levels are relatively elevated in the KOs) (Extended Data Fig. 10b).

Importantly, host gene regulation is directly dependent on the presence of the intronic L1, as deletion of select MORC2/HUSH-bound L1s from the intron led to the upregulation of host mRNA to a level commensurate with the magnitude of changes caused by HUSH/MORC2 KO (Fig. 4d,e and Extended Data Fig. 10c,d). Thus, dampening expression levels of an active gene can be a by-product of a retrotransposition event and associated HUSH/MORC2-mediated L1 silencing (Fig. 4f). Although observed effects on active host genes are only modulatory, they occur to various extents at hundreds of human genes, illustrating how TE activity can rewire host gene expression patterns.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 May; accepted 28 November 2017.

Published online 6 December 2017.

- Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
- Levin, H. L. & Moran, J. V. Dynamic interactions between transposable elements and their hosts. *Nat. Rev. Genet.* **12**, 615–627 (2011).
- Beck, C. R., Garcia-Perez, J. L., Badge, R. M. & Moran, J. V. LINE-1 Elements in Structural Variation and Disease. *Annu. Rev. Genomics Hum. Genet.* **12**, 187–215 (2011).
- Mita, P. & Boeke, J. D. How retrotransposons shape genome regulation. *Curr. Opin. Genet. Dev.* **37**, 90–100 (2016).
- Goodier, J. L. Restricting retrotransposons: a review. *Mob. DNA* **7**, 16 (2016).
- Chuong, E. B., Elde, N. C. & Feschotte, C. Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.* **18**, 71–86 (2017).
- Philippe, C. *et al.* Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci. *eLife* **5**, e13926 (2016).
- Tchassovnikarova, I. A. *et al.* Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. *Science* **348**, 1481–1485 (2015).
- Moran, J. V. *et al.* High Frequency Retrotransposition in Cultured Mammalian Cells. *Cell* **87**, 917–927 (1996).
- Morgens, D. W. *et al.* Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. *Nat. Commun.* **8**, 15178 (2017).
- Morgens, D. W., Deans, R. M., Li, A. & Bassik, M. C. Systematic comparison of CRISPR-Cas9 and RNAi screens for essential genes. *Nat. Biotechnol.* **34**, 634–636 (2016).
- Suzuki, J. *et al.* Genetic Evidence That the Non-Homologous End-Joining Repair Pathway Is Involved in LINE Retrotransposition. *PLoS Genet.* **5**, (2009).
- Chance, P. F. *et al.* Linkage of the gene for an autosomal dominant form of juvenile amyotrophic lateral sclerosis to chromosome 9q34. *Am. J. Hum. Genet.* **62**, 633–640 (1998).
- Németh, A. H. *et al.* Autosomal Recessive Cerebellar Ataxia with Oculomotor Apraxia (Ataxia-Telangiectasia-Like Syndrome) Is Linked to Chromosome 9q34. *Am. J. Hum. Genet.* **67**, 1320–1326 (2000).
- Albulym, O. M. *et al.* MORC2 mutations cause axonal Charcot-Marie-Tooth disease with pyramidal signs. *Ann. Neurol.* **79**, 419–427 (2016).
- Schottmann, G., Wagner, C., Seifert, F., Stenzel, W. & Schuelke, M. MORC2 mutation causes severe spinal muscular atrophy-phenotype, cerebellar atrophy, and diaphragmatic paralysis. *Brain* **139**, e70–e70 (2016).
- Brégnard, C. *et al.* Upregulated LINE-1 Activity in the Fanconi Anemia Cancer Susceptibility Syndrome Leads to Spontaneous Pro-inflammatory Cytokine Production. *EBioMedicine* **8**, 184–194 (2016).
- Ostertag, E. M., Luning Prak, E. T., DeBerardinis, R. J., Moran, J. V. & Kazazian, H. H. Determination of L1 retrotransposition kinetics in cultured cells. *Nucleic Acids Res.* **28**, 1418–1423 (2000).
- Han, J. S. & Boeke, J. D. A highly active synthetic mammalian retrotransposon. *Nature* **429**, 314–318 (2004).
- Wagstaff, B. J., Barner-Joi, M. & Roy-Engel, A. M. Evolutionary Conservation of the Functional Modularity of Primate and Murine LINE-1 Elements. *PLOS ONE* **6**, e19672 (2011).
- Tchassovnikarova, I. A. *et al.* Hyperactivation of HUSH complex function by Charcot-Marie-Tooth disease mutation in MORC2. *Nat. Genet.* **49**, 1035–1044 (2017).
- Moisiard, G. *et al.* MORC Family ATPases Required for Heterochromatin Condensation and Gene Silencing. *Science* **336**, 1448–1451 (2012).
- Pastor, W. A. *et al.* MORC1 represses transposable elements in the mouse male germline. *Nat. Commun.* **5**, 5795 (2014).
- Garcia-Perez, J. L. *et al.* LINE-1 retrotransposition in human embryonic stem cells. *Hum. Mol. Genet.* **16**, 1569–1577 (2007).
- Han, J. S., Szak, S. T. & Boeke, J. D. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* **429**, 268–274 (2004).
- Saint-André, V., Batsché, E., Rachez, C. & Muchardt, C. Histone H3 lysine 9 trimethylation and HP1 γ favor inclusion of alternative exons. *Nat. Struct. Mol. Biol.* **18**, 337–344 (2011).
- Khan, H., Smit, A. & Boissinot, S. Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. *Genome Res.* **16**, 78–87 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Moran for the LRE-GFP plasmid and Astrid Engel for the codon-optimized L1 construct. We thank D. Fuentes, A. Spencley, R. Srinivasan, J. Mohammed, V. Bajpai, K. Tsui, G. Hess, D. Morgens, G. Cornelis for assistance and discussions. We thank K. Cimprich, A. Fire, A. Urban for comments on the manuscript. This work was funded by Jane Coffin Childs Memorial Fund for Medical Research (N.L.), NSF DGE-114747 (C.H.L.), NIH R01HG008150 (M.C.B.), NIH 1DP2HD084069-01 (M.C.B.), NIH R01 GM112720 and Howard Hughes Medical Institute grants (J.W.).

Author Contributions N.L., C.L., T.S., J.W., M.B. designed and performed experiments, analyzed data and wrote the manuscript. E.G., C.L., J.W., M.B. initiated the K562 genome-wide screen. B.G. analyzed smFISH data. J.W., M.B. supervised the entire work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to J.W. (wyssocka@stanford.edu) or M.B. (bassik@stanford.edu).

Reviewer Information Nature thanks D. Bouché and the other anonymous reviewer(s) for their contribution to the peer review of this work.

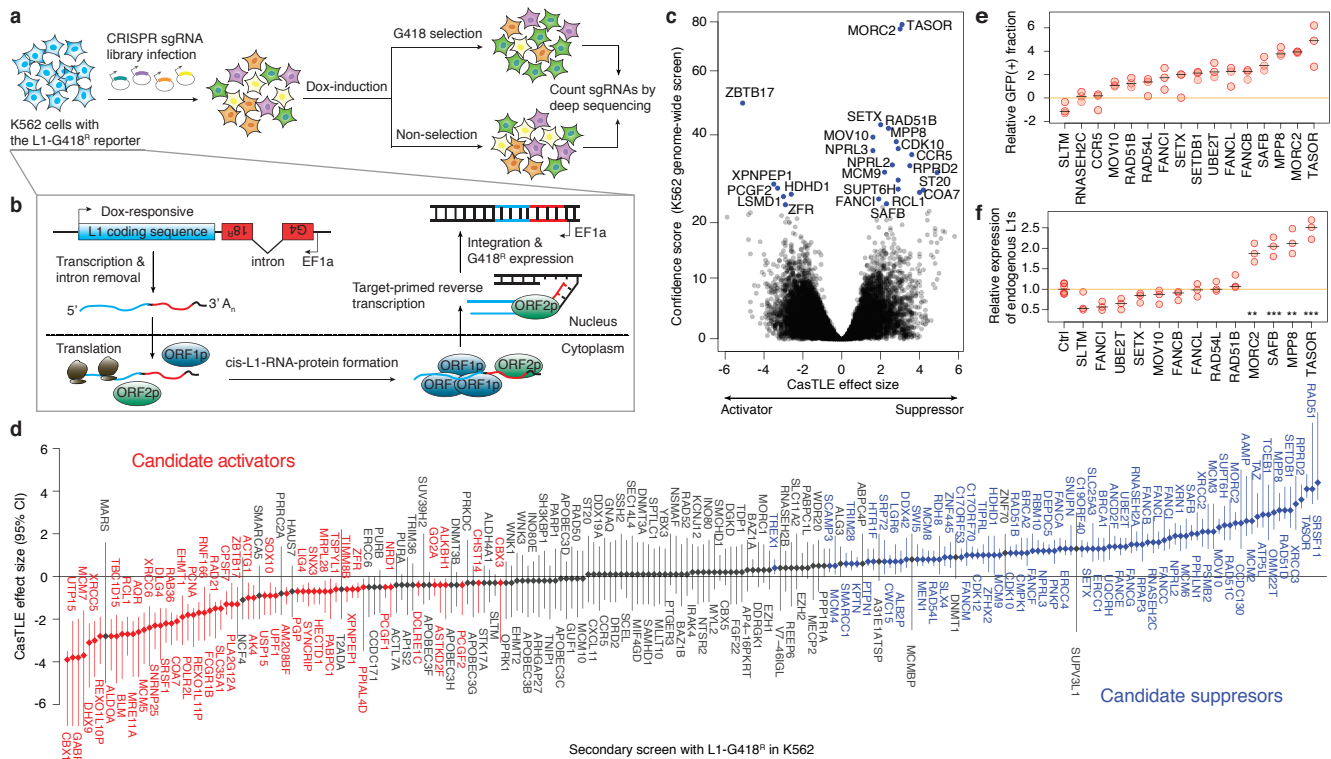


Figure 1 | Genome-wide screen for L1 activators and suppressors in K562 cells. **a**, Schematic for the screen. **b**, Schematic for the L1-G418^R retrotransposition. **c**, CasTLE analysis of (n = 2) independent K562 genome-wide screens. Genes at 10% FDR cutoff colored in blue, CasTLE likelihood ratio test¹¹. **d**, The maximum effect size (center value) estimated by CasTLE from two independent K562 secondary screens with 10 independent sgRNAs per gene. Bars, 95% credible interval (CI). L1 activators, red; L1 suppressors, blue; insignificant genes whose CI

include 0, gray. **e**, L1-GFP retrotransposition in control (infected with negative control sgRNAs, hereinafter referred to as 'Ctrl') and mutant K562 cells as indicated. GFP(+) cell fractions normalized to Ctrl. Center value as median. n = 3 biological replicates per gene. **f**, RT-qPCR measuring endogenous L1s expression in mutant K562 cells, normalized to Ctrl. Center value as median. n = 3 technical replicates per gene. **P < 0.01; ***P < 0.001; two-sided Welch t-test.

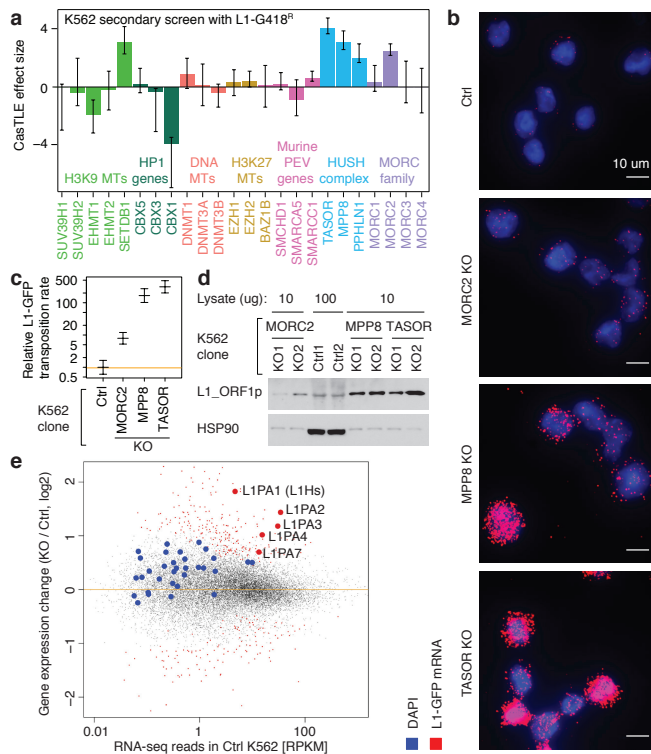


Figure 2 | HUSH and MORC2 silence L1 transcription to inhibit retrotransposition. **a.** The maximum effect size (center value) of indicated heterochromatin regulators, estimated by CasTLE from two independent K562 secondary screens with 10 independent sgRNAs per gene. Error bars, 95% credible intervals. **b.** Visualization of L1-GFP mRNAs in dox-induced K562 clones, from single smFISH experiment that was independently repeated twice with similar results. See also Extended Data Fig. 4d,e. **c.** L1-GFP retrotransposition rate¹⁸ (center value) in K562 clones, from logistic regression fit of the GFP(+) cell counts at 7 time points (0, 5, 10, 15, 20, 25, 30 days post-induction) and two independent clones per gene. Over 200 GFP(+) cells per cell count. Data normalized to Ctrl. Bar, 95% credible interval. **d.** Endogenous L1_ORF1p level in K562 clones by western blots, HSP90 as loading control. Three experiments repeated independently with similar results. **e.** RNA-seq read counts from MORC2 KO, MPP8 KO and TASOR KO K562 clones, compared to Ctrl RNA-seq reads. $n = 6 + 2$ biologically independent RNA-seq experiments). Dots represent transcripts; large dots represent L1 transcripts. Red, significant changes (padj < 0.1, DESeq analysis); blue and gray, insignificant changes.

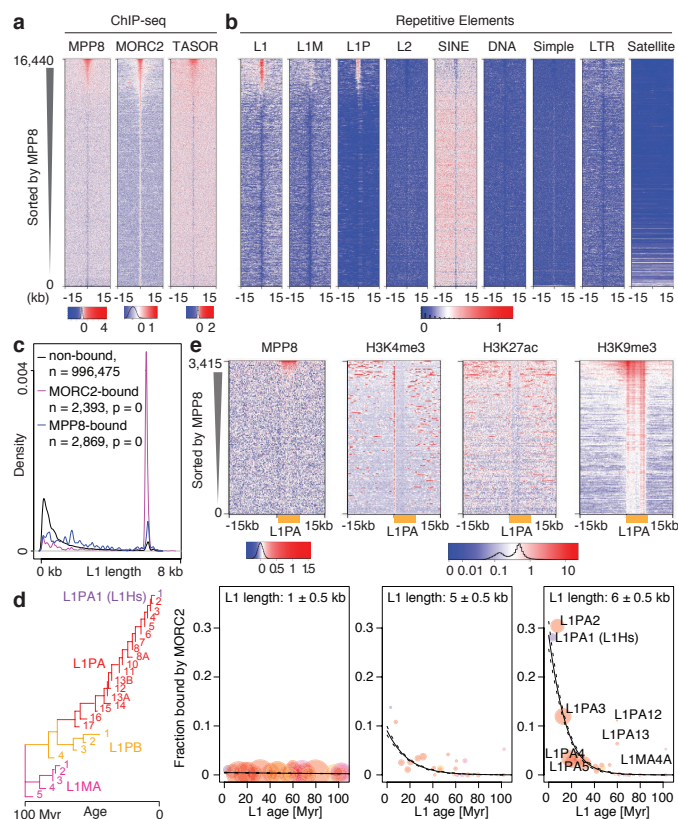


Figure 3 | HUSH/MORC2 target young full-length L1s in euchromatic environment. **a.** Heatmaps showing signal enrichment of ChIPs with indicated antibodies in K562 cells, sorted by MPP8 ChIP signal and centered on MPP8 and MORC2 peaks. Plotted is normalized ChIP signal (Ctrl subtracted with corresponding KO). **b.** Heatmaps showing MPP8 and MORC2 ChIP signal enrichment over repetitive elements, centered and sorted as in (a). **c.** Size distribution of the L1s bound or unbound by MORC2 or MPP8 in K562 cells. P-values, two-tailed Kolmogorov-Smirnov test. **d.** Fraction of MORC2-bound L1s (center values) as function of L1 length (three size classes are presented) and age (predicted from the phylogenetic analysis²⁷) in K562 cells. Colored circles represent L1 families, with areas proportional to count of L1 instances with indicated age and length. $n = 1,501$ MORC2-bound L1 + 200,160 unbound L1. $p = 2.2 \times 10^{-90}$ for age-length interaction term, lower for simple terms (ANOVA, χ^2 test), plotted logistic regression lines with 95% credible interval. **e.** Heatmaps showing signal enrichment of ChIPs with indicated antibodies in K562 cells, centered on the 5' end of full-length L1PAs.

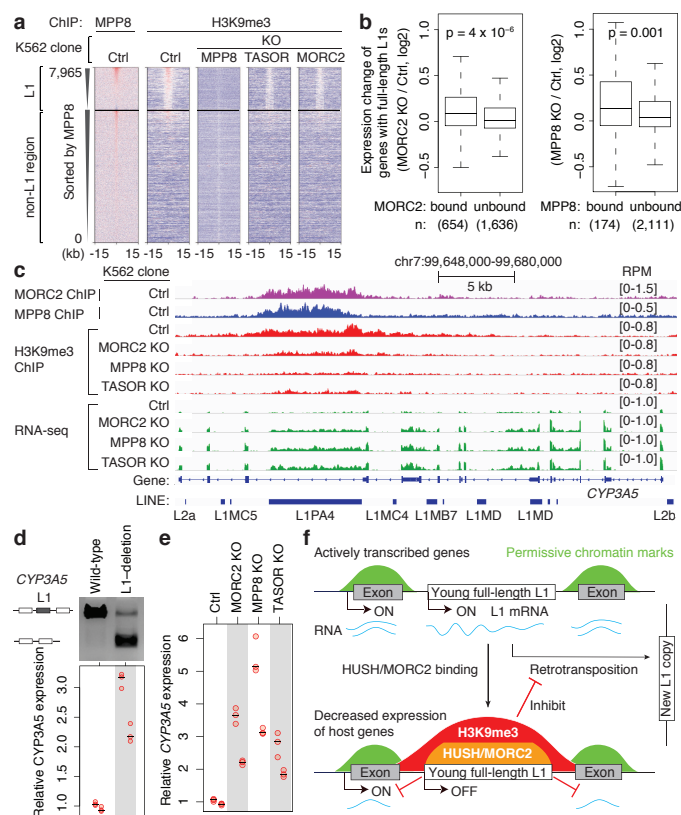


Figure 4 | HUSH/MORC2 binding at L1s decreases active host gene expression. **a**, Heatmaps showing MPP8 and H3K9me3 ChIP signal enrichment, centered on MPP8 and MORC2 summits and separated by L1 presence or absence. **b**, Expression change of genes with intronic full-length L1s that are bound or unbound by MORC2 or MPP8 (RNA-seq reads from KO K562 clones compared to Ctrl). Box plots show median and interquartile range (IQR), whiskers are $1.5 \times$ IQR. p -value, two-sided Mann-Whitney-Wilcoxon test. **c**, Genome browser tracks: HUSH/MORC2 loss causing H3K9me3 decrease at the target L1 and expression increase at both the target L1 and its host gene, independently repeated once with similar results. **d**, Deleting the target intronic L1 from *CYP3A5* in K562 increases *CYP3A5* expression, by RT-qPCR normalized to wild-type sample. $n = 2$ biological replicates \times 3 technical replicates (center value as median). Gel image confirms L1 deletion; two experiments repeated independently with similar results. **e**, RT-qPCR for *CYP3A5* expression in K562 clones, normalized to Ctrl. $n = 2$ biological replicates \times 3 technical replicates (center value as median). **f**, Model: HUSH/MORC2 bind young full-length L1s within transcriptionally active genes, and promote H3K9me3 deposition at target L1s to silence L1 transcription. This pathway not only inhibits L1 retrotransposition, but also decreases host gene expression.

METHODS

Cell culture and antibodies. K562 cells (ATCC) were grown in Roswell Park Memorial Institute (RPMI) 1640 Medium (11875093, Life Technologies) supplemented with 10% Fetal Bovine Serum (Fisher, Cat# SH30910), 2 mM L-glutamine (Fisher, Cat# SH3003401) and 1% penicillin-streptomycin (Fisher, Cat#SV30010), and cultured at 37°C with 5% CO₂. HeLa cells (ATCC) were grown in Dulbecco's Modified Eagle's Medium (Life Technologies, Cat# 11995073) supplemented with 10% FBS, 2 mM L-glutamine, and 1% penicillin-streptomycin, and cultured at 37°C with 5% CO₂. H9 human ES cells were expanded in feeder-free, serum-free medium mTeSR-1 from StemCell technologies, passaged 1:6 every 5–6 days using accutase (Invitrogen) and re-plated on tissue culture dishes coated overnight with growth-factor-reduced matrigel (BD Biosciences). Male mouse embryonic stem cells (R1) were grown as described²⁸. Cell cultures were routinely tested and found negative for mycoplasma infection (MycAlert, Lonza).

Rabbit MORC2 antibody (A300-149A, Bethyl Laboratories), Rabbit MPP8 antibody (16796-1-AP, Protein Technologies Inc), Rabbit TASOR antibody (HPA006735, Atlas Antibodies) were used in Western blots (1:1000 dilution) and ChIP assays. Mouse anti-LINE-1 ORF1p antibody (MABC1152, Millipore)²⁹, Rabbit HSP90 (C45G5, Cell Signalling, #4877), Beta actin antibody (ab49900, Abcam) were used in Western blots. Histone H3 (tri-methyl K9) antibody (ab8898, Abcam) and RNA Pol II (Santa Cruz Biotechnology, N-20 sc-899) were used in ChIP assays.

L1 reporters. The L1-ORF1-ORF2 sequence is derived from the LRE-GFP³⁰, a gift from John Moran. To make the L1-GFP reporter, we used Gibson assembly to clone the L1_ORF1/2 fragment and a GFP-B-globin-intron cassette driven by the mammalian promoter EF1a into the pB transgene using a dox inducible promoter (modified from PBQM812A-1, System Biosciences) to drive the L1 sequence and a UBC-RTTA3-ires Blast as a selectable marker for reporter integration. To make the L1-G418^R reporter, we replaced the GFP-B-globin-intron fragment in the L1-GFP reporter with a NEO-intron-NEO cassette driven by the mammalian promoter EF1a. The codon-optimized L1-ORF1-ORF2 sequence in our (opt)-L1 reporter is derived from the SynL1_optORF1_neo, a gift from Astrid Engel³¹. We replaced the self-splicing Tetrahymena NEO-intron-NEO cassette with the neo-B-globin-intron-neo cassette driven by the EF1a promoter or the GFP-B-globin-intron-GFP cassette driven by the EF1a promoter. This L1-syn-ORF1-ORF2-indicator cassette was inserted into the pB transgene using a dox inducible promoter and a UBC-RTTA3-ires Blast, as described above.

Genome-wide screen in K562 cells. The K562 cell line (with a BFP-Cas9 lentiviral transgene) was nucleofected with the pB-tetO-L1-G418^R/Blast construct and the piggyBac transposase (PB210PA-1, System Biosciences) following the manufacturer's instructions (Lonza 2b nucleofector, T-016 program). The nucleofected cells were sorted using limiting dilution in 96-well plates, and positive clones were screened first for sensitivity to Blast, and then the ability to generate G418 resistant cells after dox induction. The Cas9/L1-G418^R cells were lentivirally infected with a genome-wide sgRNA library as described¹⁰, containing ~200,000 sgRNAs targeting 20,549 protein-coding genes and 13,500 negative control sgRNAs at an MOI of 0.3–0.4 (as measured by the mCherry fluorescence from the lentiviral vector), and selected for lentiviral integration using puromycin (1 µg/ml) for 3 days as the cultures were expanded for the screens. In duplicate, 200x10⁶ library-infected cells were dox-induced (1 µg/ml) for 10 consecutive days, with a logarithmic growth (500k cells/ml) maintained each day of the dox-induction. After dox-induction, the cells were recovered in normal RPMI complete media for 24 hours, and then split into the G418-selection condition (300 µg/ml G418, Life Technologies, Cat# 11811031) and non-selection conditions. After 7 days of maintaining cells at 500k/ml, 200 M cells under each condition were recovered in normal RPMI media for 24 hours, before they were pelleted by centrifugation for genomic DNA extraction using Qiagen DNA Blood Maxi kit (Cat# 51194) as described³². The sgRNA-encoding constructs were PCR-amplified using Agilent Herculase II Fusion DNA Polymerase (Cat# 600675) (See Table S4 for the primer sequences used). These libraries were then sequenced across two Illumina NextSeq flow cells (~40 M reads per condition; ~200x coverage per library element). Computational analysis of genome-wide screen was performed as previously described^{10,11} using CasTLE, which is a maximum likelihood estimator that uses a background of negative control sgRNAs as a null model to estimate gene effect sizes. See Table S1 for the K562 genome-wide screen results.

Secondary screen in K562 cells. The secondary screen library included the following, non-comprehensive sets of genes (253 genes in total, ~10 sgRNAs per gene, plus 2500 negative control sgRNAs): all genes falling within ~30% FDR from the K562 genome-wide screen (~150 genes), genes known to be functionally related to the 30% FDR genes, genes previously implicated in L1 biology, and genes involved in epigenetic regulation or position effect variegation (see Table S2 for a complete list). The library oligos were synthesized by Agilent Technologies and cloned into pMCB320 using BstXI/BlpI overhangs after PCR amplification. The Cas9/L1-G418^R (or Cas9/(opt)-L1-G418^R) K562 cell line was lentivirally infected with the secondary library (~4,500 elements) at an MOI of 0.3–0.4 as described

previously³³. After puromycin selection (1 µg/ml for 3 days) and expansion, 40 M (~9,000 coverage per library element) cells were dox-induced for 10 days in replicate, recovered for 1 day, and split for 7-day G418-selection and non-selection conditions, with a logarithmic growth (500k cells/ml) maintained as in the K562 genome-wide screen. 10M cells under each condition were used for genomic extractions, sequenced (~6–10M reads per condition; ~1000–2000x coverage per library element) and analyzed using CasTLE as described above^{10,11}. See Table S2 for the K562 secondary screen results with L1-G418^R and (opt)-L1-G418^R.

Genome-wide screen and Secondary screen in HeLa cells. The pB-tetO-L1-G418^R/Blast construct was integrated into Cas9 expressing HeLa cells with piggyBac transposase via nucleofection (Lonza 2b nucleofector, I-013 program) following the manufacturer's instructions. The Cas9/L1-G418^R HeLa cells were blasticidin (10 µg/ml) selected, screened for sensitivity to G418 and the ability to generate G418 resistance cells after dox induction, and lentivirally infected with the genome-wide sgRNA library or with the secondary sgRNA library. Infected cells were then puromycin selected (1 µg/ml) for 5 days and expanded for the screens.

For the genome-wide screen, ~200x10⁶ Cas9/L1-G418^R HeLa cells (~1,000x coverage of sgRNA library) were dox-induced for 10 days in replicate, recovered for 1 day, and split for 8-day G418-selection and non-selection conditions, with cells being split every other day to maintain the sgRNA library at a minimum of ~350x coverage. ~200M (1,000x coverage) cells per condition were used for genomic extractions and sequencing as described above for the K562 screens. See Table S1 for the HeLa genome-wide screen results.

For the secondary screen, ~1x10⁷ Cas9/L1-G418^R HeLa cells (~2,000x coverage of sgRNA library) were dox-induced for 10 days in replicate, recovered for 1 day, and split for 8-day G418-selection and non-selection conditions, with cells being split every other day to maintain ~400x coverage. ~5 million (1,000x coverage) cells per condition were used for genomic extractions and sequencing as described above. See Table S2 for the HeLa secondary screen results.

Validation of individual candidates using the L1-GFP retrotransposition assay. To validate the genome-wide screen hits, we infected clonal Cas9/L1-GFP K562 cells with individual sgRNAs as previously described³², 3 independent mutant cell lines per gene, each with a different sgRNA (cloned into pMCB320 using BstXI/BlpI overhangs; mU6:sgRNA; EF1a:Puromycin-t2a-mCherry). See Table S3 for sgRNA sequences. The infected cells were selected against puromycin (1 µg/ml) for 3 days, recovered in fresh RPMI medium for 1 day, and dox-induced for 10 days. Then, the percentage of GFP(+) cells was measured on a BD Accuri C6 Flow Cytometer (GFP fluorescence detected in FL1 using 488 nm laser) after gating for live mCherry(+) cells.

CRISPR-mediated deletion of individual genes and intronic L1s. To delete genes in H9 ESCs, we cloned target sgRNAs in pSpCas9(BB)-2A-GFP (PX458) as described³⁴. The sgRNA plasmids were prepared with the Nucleosin plasmid kit (Macherey Nagel) and transfected into H9 ESCs using Fugene following the manufacturer's instructions. After 48–72 hrs, GFP-positive transfected cells were sorted and expanded. Gene depletion effects were validated by western blots.

To delete the L1 from the host gene intron, we designed sgRNAs targeting both upstream and downstream side of the L1 within the intron; one was cloned into pSpCas9(BB)-2A-BFP, while the other into pSpCas9(BB)-2A-GFP. The two sgRNA plasmids were mixed at 1:1 ratio and nucleofected into K562 cells via electroporation following the manufacturer's instructions. After 48–72 hours, BFP/GFP-positive transfected cells were single-cell sorted and expanded. The genetic deletion effects were validated by PCR assay.

Western blotting. Live cells were lysed for 30 min at 4°C in protein extraction buffer (300 mM NaCl, 100 mM Tris pH 8, 0.2 mM EDTA, 0.1% NP40, 10% glycerol) with protease inhibitors and centrifuged to collect the supernatant lysate. The cell lysate was measured with Bradford reagent (Biorad), separated on SDS-PAGE gels and transferred to nitrocellulose membranes. The L1-reporter containing K562 cells had not been dox-induced when used for western blot assays characterizing endogenous L1_ORF1p levels (Fig. 2d and Extended Data Fig. 4k).

PCR and gel electrophoresis. PCR experiments characterizing the L1-G418^R retrotransposition and the deletion of intronic L1s were performed with Phusion High-Fidelity DNA Polymerase (M0530S, NEB), following the manufacturer's instructions. In general, 30 cycles of PCR reactions were performed at an annealing temperature 5°C below the T_m of the primer. No 'spliced' PCR products can be detected without dox-induction, even with 40 PCR cycles. PCR reaction products were separated on 1% agarose gels with ethidium bromide. Primer sequences are in Table S4.

qRT-PCR and PspGI-assisted qPCR. Total RNA was isolated from live cells using the RNeasy kit (74104, Qiagen) and treated with RNase-Free DNase Set (79254, Qiagen) to remove genomic DNA, according to the manufacturer's instructions. 500 ng total RNA was reverse transcribed with SuperScriptA III First-Strand Synthesis System (18080051, Life Technologies) following the manufacturer's instructions. Beta-actin mRNA was used as internal control within each RNA sample (Figs. 1f and 4d,e). The sequences of PCR primers, including the one targeting the 5'UTR of L1Hs^{35–37}, are summarized in Table S4.

Genomic DNA was isolated using PureLink Genomic DNA Mini Kit (K182001, Life Technologies) with RNase A digestion to remove contaminant RNA, according to the manufacturer's instructions. 300 ng genomic DNA per sample was digested with 50 units PspGI (R0611S, New England Biolabs) in 1x smart buffer (NEB) at 75°C for 1 hr, to cut uniquely at the intron of the GFP cassette. The reaction mixture was then used in qPCR experiments with primers flanking the intron in the GFP cassette (Table S4). Due to the PspGI digestion, the original unspliced L1-GFP reporter will not be amplified by PCR. Only newly integrated GFP cassettes, where the intron was removed during the retrotransposition process, can be PCR amplified. qPCR runs and analysis were performed on the Light Cycler 480II machine (Roche).

Northern Blotting. Northern blotting was conducted as previously described³⁸. Briefly, 15 µg of total RNA from K562 cells or H9 ESC cells was separated on the 0.7% formaldehyde agarose gel, capillary transferred overnight in 20x SSC to the Hybond N membrane (GE Healthcare), crosslinked with a Stratilinker (Stratagene), and hybridized with ³²P-labeled single-stranded DNA probes (10⁶ cpm/ml) in ULTRAhyb-Oligo Hybridization Buffer (AM8663, Life Technologies) following the manufacturer's instructions. Blots were washed two times with wash buffer (2X SSC, 0.5% SDS), and then exposed to film overnight to several days at -80°C with an intensifying screen. The sequence of oligonucleotide probes is in Table S3.

Single molecule FISH. Single molecule FISH (smFISH) assays were performed following the Affymetrix Quantigene ViewRNA ISH Cell Assay user manual. 2.5–5 million live K562 cells were fixed within 4% formaldehyde in 1x PBS for 60 mins at RT, resuspended in 1x PBS, pipetted onto poly-L-lysine coated glass cover slip (~20,000 total cells/spot; spread out with a pipette tip), and baked in dry oven at 50±1°C for 30 minutes to fix the cells onto the glass slip, followed by digestion with Protease QS (1:4000) in 1x PBS for 10 minutes at RT. Cells were hybridized with smFISH probes, designed to target beta actin mRNA (FITC channel) and the L1-GFP reporter mRNA (Cy3 channel), DAPI stained for 5 mins, and mounted with Prolong Gold Antifade Reagent (10 ml/sample). Images were taken by spinning disk confocal microscope equipped with 60x 1.27NA water immersion objective with an effective pixel size of 108x108 nm. Specifically, for each field of view, a z-series of 8 µm is taken with 0.5 µm/z-step for all 3 channels. For quantitation, maximum-projected images from the z-series is used and analyzed by a custom-written matlab script. In brief, all images are first subtracted with the background determined with the OTSU method³⁹ from the log-transformed image after pillbox blurring with a radius of 3 pixels. mRNA puncta are segmented by tophat filter using the background subtracted images and only the ones above 25th percentile intensity of all segmented puncta are taken for downstream analysis. Each punctum is then assigned to the nuclear mask identified by image areas above the previously determined background. For each single cell, the assigned pixel area of L1-GFP mRNA is then normalized to the assigned pixel area of beta-actin mRNA per cell.

RNA-seq. Two independent biological replicates of K562 cells in culture were extracted to isolate DNA-free total RNA sample, using the RNeasy kit (74104, Qiagen) combined with the RNase-Free DNase Set (79254, Qiagen). PolyA-selected RNA were isolated using 'Dynabeads mRNA Purification Kit for mRNA Purification from Total RNA preps' (610-06, Life Technologies) following the manuals. 100 ng polyA-selected RNA was fragmented with NEBNext Magnesium RNA Fragmentation Module (E6150S, New England Biolabs), and used for first strand cDNA synthesis with SuperScriptII (18064-014, Invitrogen) and random hexamers, followed by second strand cDNA synthesis with RNaseH (18021-014, Invitrogen) and DNA PolI (18010-025, Invitrogen). The cDNA was purified, quantified, multiplexed and sequenced with 2x 75bp pair-end reads on an Illumina NEXT-seq (Stanford Functional Genomics Facility).

RNA-seq reads were aligned to hg38 reference genome with hisat2 (--no-mixed, --no-discordant) without constraining to known transcriptome. Known (genome 25) and de-novo transcript coverages were quantified with featureCount. Repeat Masker coverage was quantified with bedtools coverage. Reads mapping to the same repeat family were then tabulated together, since individual read coverage was too low to obtain meaningful results. Differential expression analysis of join gene-repeat data was performed with DESeq2⁴⁰.

ChIP-seq. Two replicates of ChIP experiments per sample were performed as previously described^{41,42}. Approximately 0.5–1 × 10⁷ cells in culture per sample were crosslinked with 1% paraformaldehyde (PFA) for 10 min at room temperature (RT), and quenched by 0.125 M glycine for 10 min at RT. Chromatin was sonicated to an average size of 0.2–0.7 kb using a Covaris (E220 evolution). Sonicated chromatin was incubated with 5–10 µg antibody bound to 100 µl protein G Dynabeads (Invitrogen) and incubated overnight at 4°C, with 5% kept as input DNA. Chromatin was eluted from Dynabeads after five times wash (50 mM Hepes, 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Na-deoxycholate), and incubated at 65°C water bath overnight (12–16 hrs) to reverse crosslinks. ChIP DNA were subject to end repair, A-tailing, adaptor ligation and cleavage with USER enzyme, followed by size selection to 250–500 bp and amplification with NEBNext sequencing primers. Libraries were purified, quantified, multiplexed (with NEBNext Multiplex

Oligos for Illumina kit, E7335S) and sequenced with 2x 75 bp pair-end reads on an Illumina NEXT-seq (Stanford Functional Genomics Facility).

ChIP-seq reads were trimmed with cutadapt (-m 50 -q 10) and aligned with bowtie2 (version 2.2.9, --no-mixed --no-discordant --end-to-end -maxins 500) to the hg38 reference genome. ChIP peaks were called with macs2 (version 2.1.1.20160309) callpeak function with broad peak option and human genome effective size using reads from corresponding loss of gene lines as background model. Visualization tracks were generated with bedtools genomecov (-bg -scale) with scaling factor being 10⁶/number aligned reads and converted to bigWig with bedGraphToBigWig (Kent tools). BigWigs were plotted with IGV browser. Individual alignments were inspected with IGB browser.

Heatmaps were generated by intersecting bam alignment files with intervals of interest (bedtools v2.25.0), followed by tabulation of the distances of the reads relative to the center of the interval and scaling to account for total aligned read numbers (10⁶/number aligned). Heatmaps were plotted using a custom R function. Aggregate plots were generated by averaging rows of the heatmap matrix. For ChIPs in Ctrl and KO K562 clones, ChIP-seq signals in the corresponding KO cells were used as the null reference.

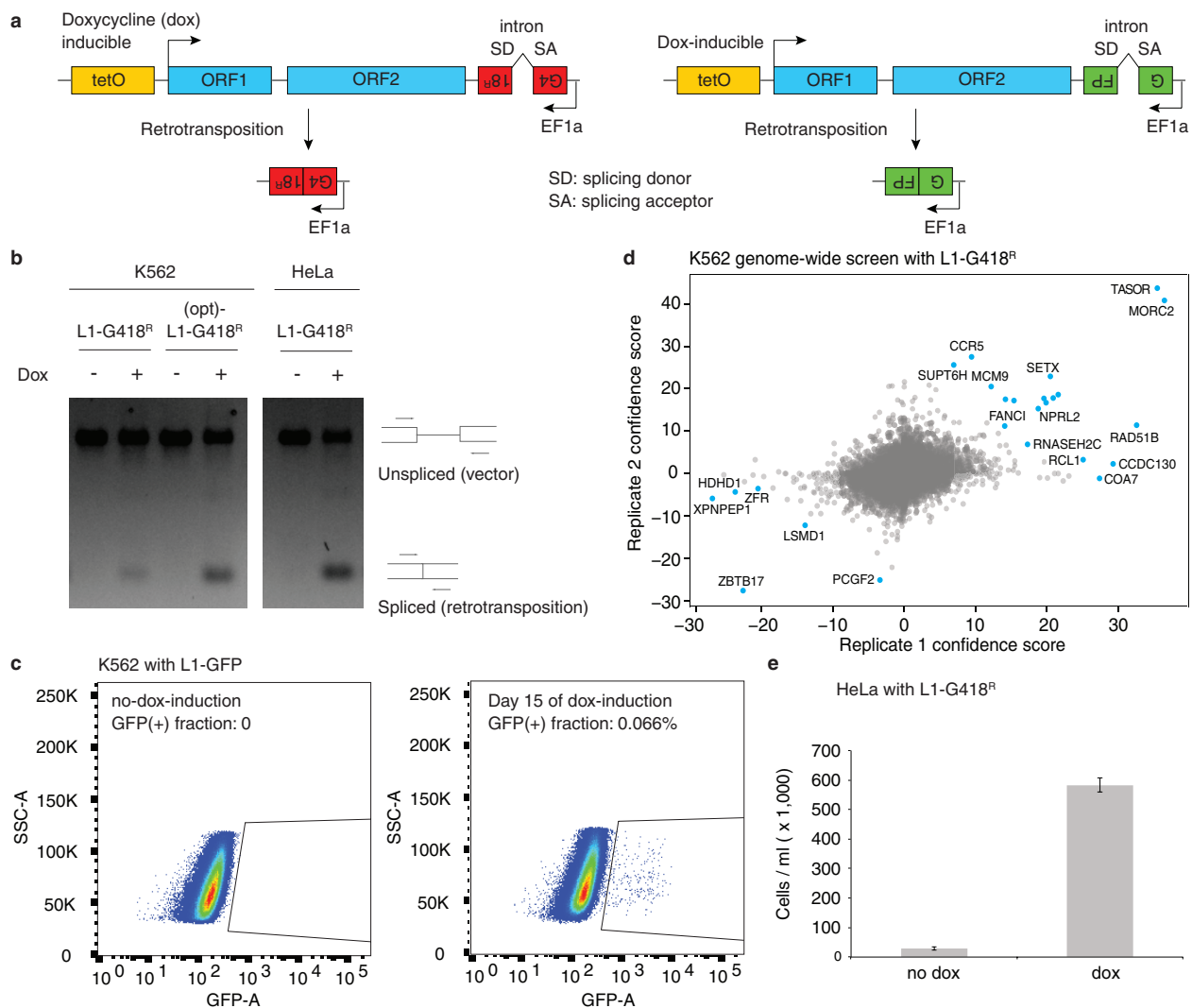
For ChIP-seq repetitive sequence relationship analysis, repeat masker was intersected with ChIP-seq peak calls to classify each masker entry as MPP8 bound, MORC2-bound or unbound. Enriched families of repeats were identified with R fisher.test() followed by FDR correction with qvalue(). Distribution of sizes of occupied vs non-occupied L1 was plotted using R density() with sizes being taken from repeat masker. ks.test() was used to reject null hypothesis that distribution of sizes for bound and unbound L1s is the same. To investigate relationship between L1 age, length and occupancy, logistic regression was performed with R glm() engine.

Quantitative analysis of H3K9me3 changes was performed by first identifying regions of significant enrichment in each sample relative to corresponding input sample (macs2 callpeak), merging the intervals into a common superset. This superset was joined with a decoy randomized set of intervals, twice the size of actual experimental interval set, with the same size distribution (bedtools shuffle). Next the read coverage was determined for each sample (bedtools coverage) and regions with significant change together with fold changes were identified using DESeq2⁴⁰. H3K9me3 regions were classified into bound vs unbound by performing intersect with MORC2 and MPP8 ChIP peak calls.

Data availability. All sequencing data generated in this work has been deposited at GEO under the accession number: GSE95374. H3K4me3 and H3K27ac K562 ChIP-seq datasets in Fig. 3e are from BioProject (accession number PRJEB8620). hESC RNA-seq datasets in Extended Data Fig. 8c are from SRA run entries SRR2043329 and SRR2043330. The complete results of genome-wide screens in K562 and HeLa cells are in Table S1; The complete results of secondary screens in K562 and HeLa cells are in Table S2. The sequences of gRNAs and oligonucleotides used in this work are in Table S3 and Table S4. The uncropped scans with size marker indications are summarized in the Supplementary Figure. All data are available from the corresponding author upon reasonable request.

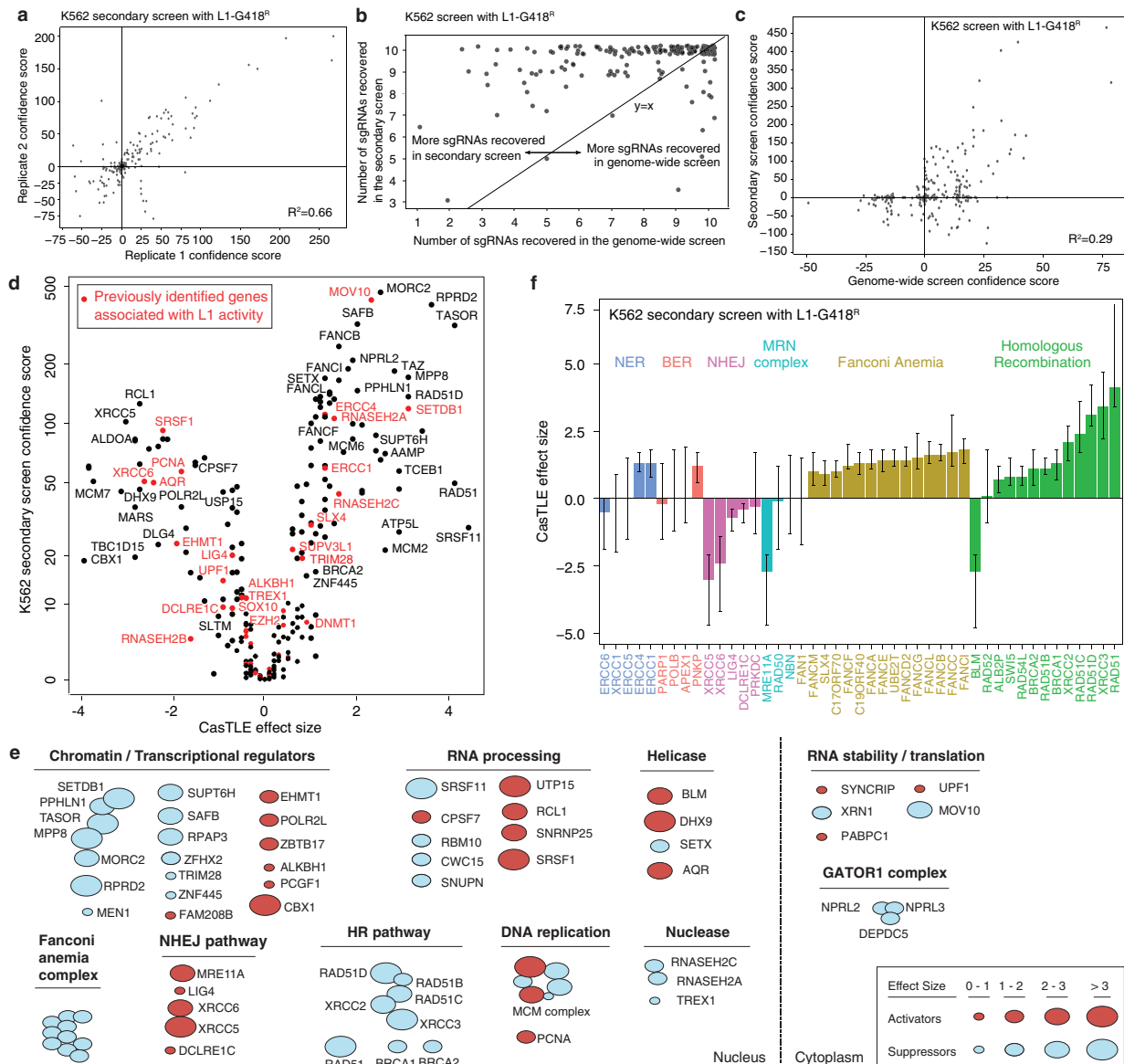
Code availability. Detailed Data and further code information are available on request from the authors.

28. Buecker, C. *et al.* Reorganization of Enhancer Patterns in Transition from Naive to Primed Pluripotency. *Cell Stem Cell* **14**, 838–853 (2014).
29. Taylor, M. S. *et al.* Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. *Cell* **155**, 1034–1048 (2013).
30. Brouha, B. *et al.* Evidence Consistent with Human L1 Retrotransposition in Maternal Meiosis I. *Am. J. Hum. Genet.* **71**, 327–336 (2002).
31. Gasior, S. L., Roy-Engel, A. M. & Deininger, P. L. ERCC1/XPF limits L1 retrotransposition. *DNA Repair* **7**, 983–989 (2008).
32. Deans, R. M. *et al.* Parallel shRNA and CRISPR-Cas9 screens enable antiviral drug target identification. *Nat. Chem. Biol.* **12**, 361–366 (2016).
33. Bassik, M. C. *et al.* A Systematic Mammalian Genetic Interaction Map Reveals Pathways Underlying Ricin Susceptibility. *Cell* **152**, 909–922 (2013).
34. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819–823 (2013).
35. Coufal, N. G. *et al.* L1 Retrotransposition in Human Neural Progenitor Cells. *Nature* **460**, 1127–1131 (2009).
36. Shukla, R. *et al.* Endogenous Retrotransposition Activates Oncogenic Pathways in Hepatocellular Carcinoma. *Cell* **153**, 101–111 (2013).
37. Carreira, P. E. *et al.* Evidence for L1-associated DNA rearrangements and negligible L1 retrotransposition in glioblastoma multiforme. *Mob. DNA* **7**, (2016).
38. Doucet, A. J., Wilusz, J. E., Miyoshi, T., Liu, Y. & Moran, J. V. A 3' Poly(A) Tract Is Required for LINE-1 Retrotransposition. *Mol. Cell* **60**, 728–741 (2015).
39. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **9**, 62–66 (1979).
40. Love, M., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
41. Bajpai, R. *et al.* CHD7 cooperates with PBAF to control multipotent neural crest formation. *Nature* **463**, 958–962 (2010).
42. Rada-Iglesias, A. *et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).



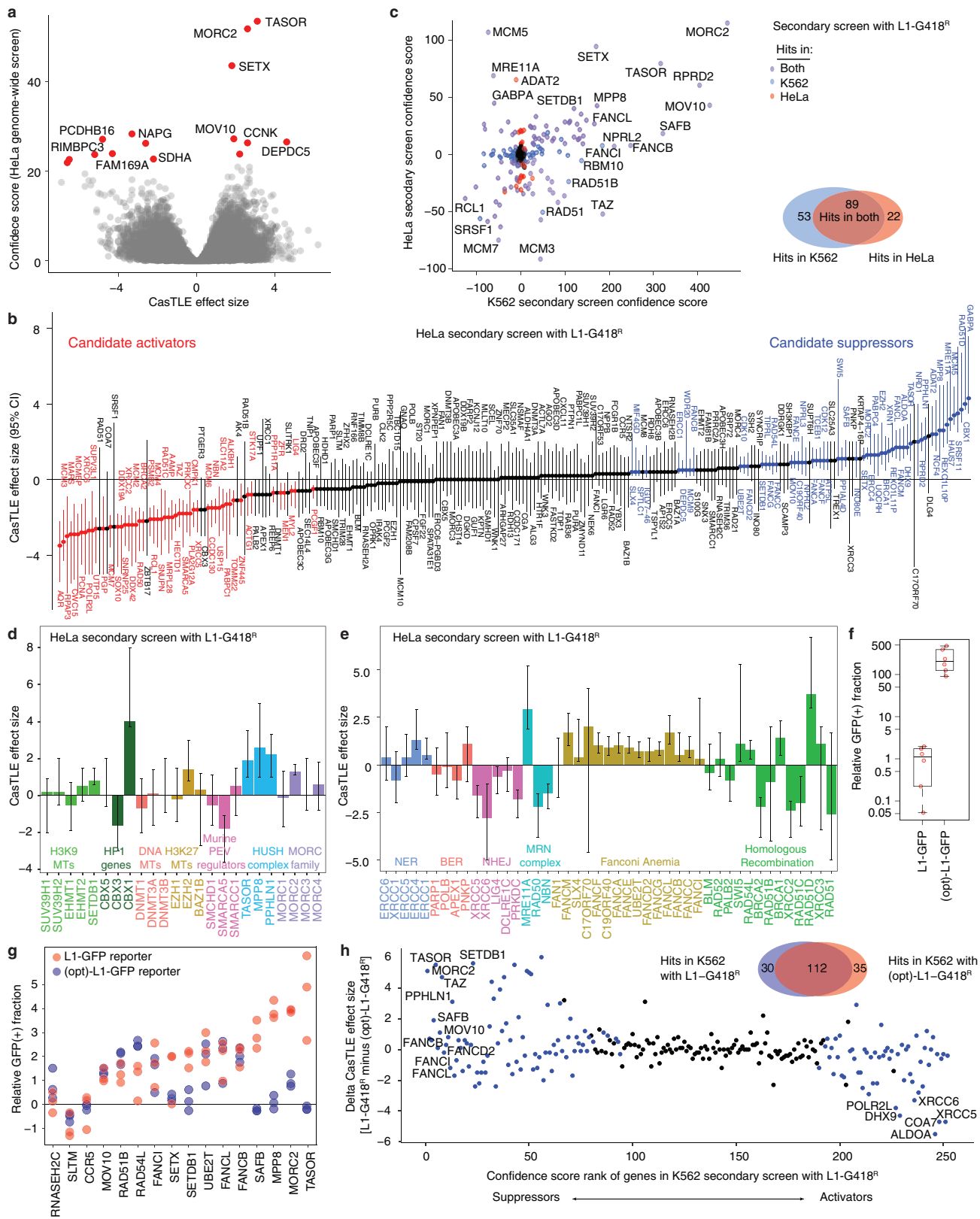
Extended Data Figure 1 | Genome-wide CRISPR/Cas9 screen for L1 regulators in K562 cells. **a.** Schematic representation of L1-G418^R and L1-GFP reporters used in this work. **b.** PCR assay on genomic DNA using primers that flank the engineered intron within the G418^R cassette. Two experiments repeated independently with similar results. The spliced PCR bands were not observed prior to dox induction in either K562 or HeLa cells, suggesting that the L1-G418^R reporter was not activated prior to the screening. However, there may exist extremely low level of reporter leakiness that is below the PCR assay detection limits. **c.** FACS results showing that the L1-GFP cells have no GFP signals without dox-induction (0 out of ~300,000 cells), and begin to produce GFP after

dox-induction. Therefore, there is insignificant level of reporter leakiness without dox-induction. Two experiments repeated independently with similar results. **d.** CasTLE analysis of genome-wide screens in K562 cells, with 20,488 genes represented as individual points. Genes falling under 10% FDR colored in blue, CasTLE likelihood ratio test¹¹. $n = 2$ biologically independent screens. **e.** HeLa with L1-G418^R are resistant to G418 after dox-induction. 7 days of dox-induction followed by 10 days of G418 selection. Live cells in equal volumes were counted in a single ($n = 1$) FACS experiment. Center value, total number of live cells. Error bar, square root of total events assuming Poisson distribution of counts.



Extended Data Figure 2 | A secondary screen identifies functionally diverse L1 regulators in K562 cells. a. Reproducibility between two independent secondary screens ($n = 2$) in K562 cells. R-squared value, linear regression model. b. The K562 secondary screen recovers more sgRNAs than the K562 genome-wide screen, suggesting a higher detection sensitivity in the secondary screen. c. Comparison of the secondary screen data (252 genes from $n = 2$ independent screens) with the genome-wide screen data ($n = 2$ independent screens) in K562 cells. R-squared

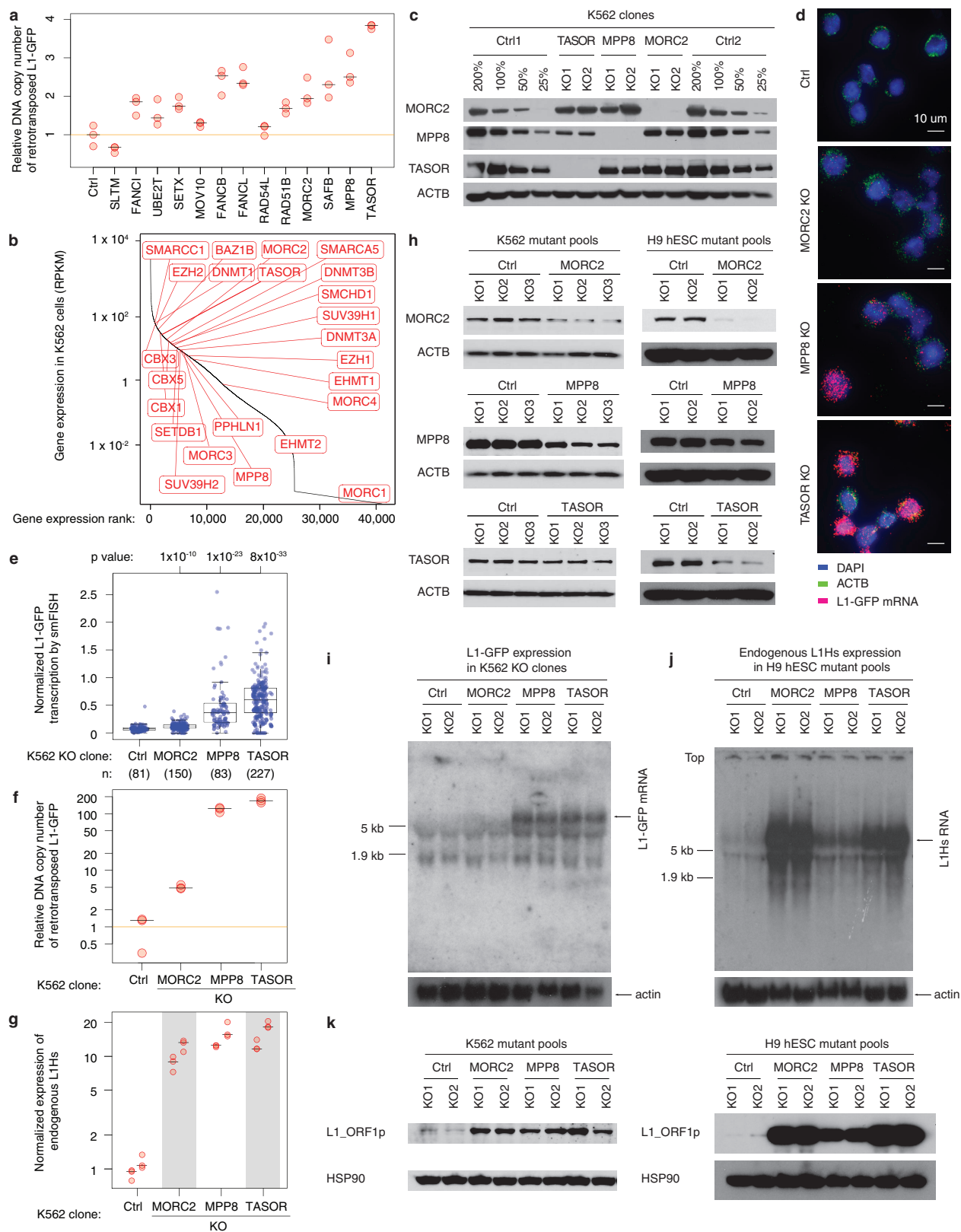
value, linear regression model. d. Volcano plot showing K562 secondary screen results (252 genes from two independent screens), with genes previously implicated in L1 biology colored in red. e. Classification diverse L1 activators and suppressors identified in K562 cells by their known biological process. f. The maximum effect size (center value) of indicated DNA repair genes, estimated by CastLE from two independent K562 secondary screens with 10 different sgRNAs per gene. Error bars, 95% credible intervals of the estimated effect size.



Extended Data Figure 3 | See next page for caption.

Extended Data Figure 3 | Screen for L1 regulators in HeLa cells and L1- sequence-dependent L1 regulators. a. CasTLE analysis of two independent genome-wide screens in HeLa cells, with 20,514 genes represented as individual points. Genes at 10% FDR cutoff colored in red, CasTLE likelihood ratio test¹¹. b. The maximum effect size (center value) estimated by CasTLE from two independent HeLa secondary screens with 10 different sgRNAs per gene. Bars, 95% credible interval (CI). L1 activators, red; L1 suppressors, blue. Genes whose CI include zero are colored in gray and are considered non-effective against L1. c. Scatter plots showing the secondary screen hits identified in K562 cells and HeLa cells (252 genes from two independent screens in each cell line), with Venn diagram comparing hits in the two cell lines is shown on the right. d. The maximum effect size (center value) of indicated heterochromatin regulators, estimated by CasTLE from two independent HeLa secondary screens with 10 different sgRNAs per gene. Error bars, 95% credible intervals of the estimated effect size. e. The maximum effect size (center value) of indicated DNA repair genes, estimated by CasTLE from two independent HeLa secondary screens with 10 different sgRNAs per gene. Error bars, 95% credible intervals of the estimated effect size.

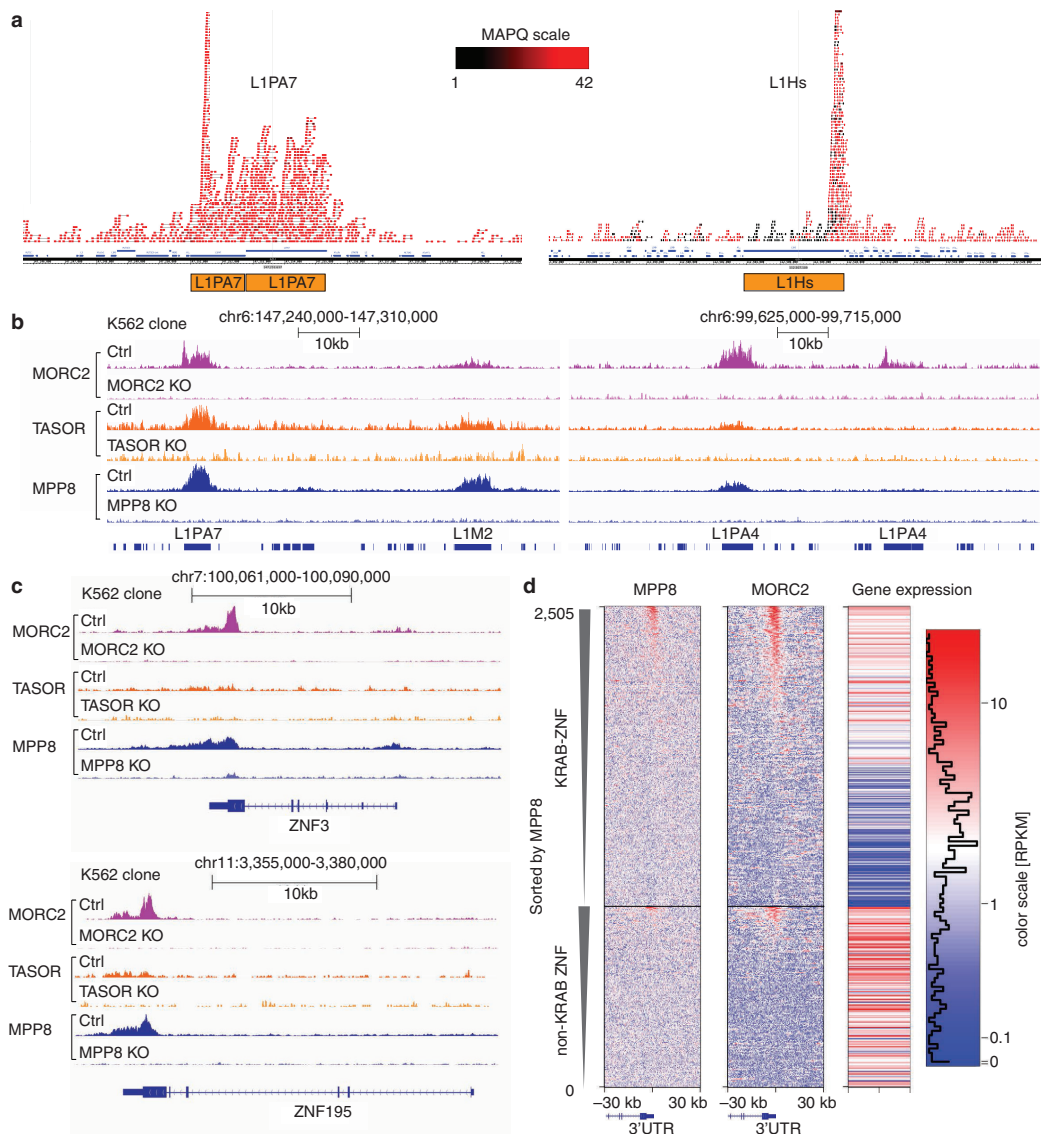
f. The (opt)-L1-GFP reporter retrotransposed more frequently than L1-GFP did in K562. The GFP(+) fraction of cells with the indicated L1 reporter after 15 days of dox induction was normalized to the L1-GFP sample. Box plots show median and interquartile range (IQR), whiskers are $1.5 \times$ IQR. $n = 6$ biologically independent replicates. g. The GFP(+) fraction of dox-induced Ctrl and mutant cell pools with the L1-GFP reporter or (opt)-L1-GFP reporter. Experiments were performed as Fig. 1e. Chromatin regulators (e.g. TASOR, MORC2, MPP8, SAFB) did not suppress the (opt)-L1-GFP reporter, in which 24% of the L1 ORF nucleotide sequence is altered, without changes in the encoded amino acid sequence^{19,20}, indicating their L1 regulation depends on the native nucleotide-sequence of L1Hs. h. K562 secondary screen with the (opt)-L1-G418^R reporter (252 genes from $n = 2$ independent screens) revealed genes that regulate retrotransposition dependent or nondependent on the native L1 nucleotide sequence. The K562 secondary screen candidates identified with L1-G418^R (252 genes from $n = 2$ independent screens) were labeled in blue. A Venn diagram comparing hits identified from the two L1-reporters is also shown.



Extended Data Figure 4 | See next page for caption.

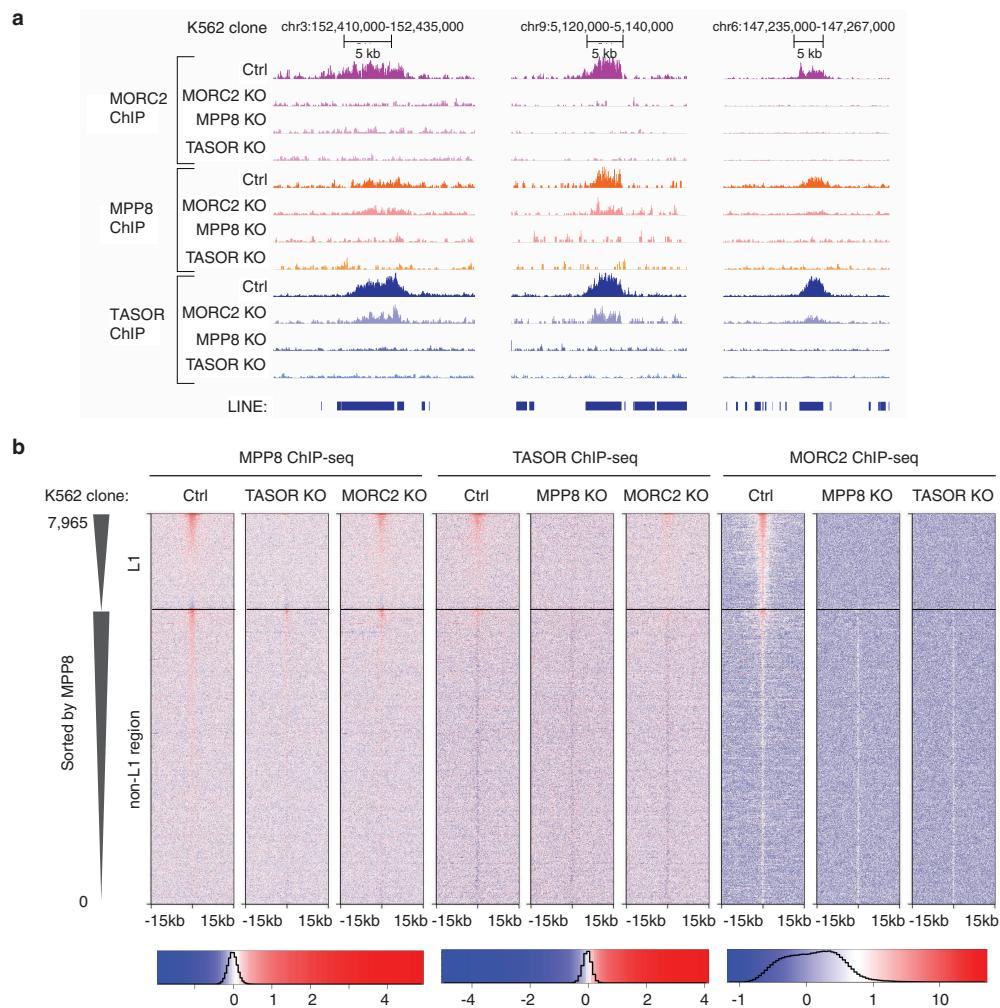
Extended Data Figure 4 | MORC2, MPP8 and TASOR silence L1 transcription. a. Relative genomic copy number of newly integrated L1-GFP reporters in the indicated mutant K562 pools after dox-induction. PspGI-assisted qPCR assay used here was designed to selectively detect spliced GFP rather than the unspliced version (see Methods section). The L1-GFP copies were normalized to beta-actin DNAs; data then normalized to Ctrl. As a putative L1 activator, SLTM shows an opposite effect on the DNA copy number, compared with L1 suppressors. Center value as median. $n = 3$ technical replicates per gene. b. RNA-seq data in Ctrl K562 cells showing that most heterochromatin regulators in Fig. 2a are expressed, supporting the selective effect of HUSH and MORC2 in L1 regulation. c. Western blots validating the knockout (KO) effects in independent KO K562 cell clones. Ctrl samples were loaded at 4 different amounts (200%, 100%, 50%, 25% of KO clones). Three experiments repeated independently with similar results. To obtain KO clones, we sorted mutant K562 pools (cells used in Fig. 1e,f) into 96-well plates, expanded cells and screened for KO clones through western blotting. Of note, all K562 KO clones were derived from the same starting L1-GFP reporter line, and thus do not differ in reporter transgene integrations among the clones. d. Representative images of single molecule FISH (smFISH) assays targeting ACTB mRNAs and RNA transcripts from L1-GFP reporters in Ctrl and KO K562 clones after 5 days of dox-induction. No signal was observed from L1-GFP reporters without dox-induction (data not shown). Two experiments repeated independently with similar results. See also panel e and Fig. 2b (showing L1-GFP mRNA only). e. Quantitation of the L1-GFP transcription level from the indicated number of K562 cells, determined by smFISH assays (panel d and Fig. 2b). The number of L1-GFP mRNA transcripts is normalized to the number of beta-actin mRNAs within each K562 cell. Box plots show median and interquartile range (IQR), whiskers are $1.5 \times$ IQR. P-value, two-sided Wilcoxon test. 95% CI for median from 1,000x bootstrap: Control: 0.059-0.082; MORC2: 0.106-0.123; MPP8: 0.264-0.410; TASOR: 0.514-0.671. f. MORC2, MPP8, and TASOR KOs increase the genomic

copy number of newly integrated L1-GFP reporters. PspGI-assisted qPCR assays were performed as in panel a), but using clonal KO K562 clones instead of mutant cell pools. Data normalized to Ctrl. $n = 3$ technical replicates, center value as median. g. MORC2 KO, MPP8 KO, and TASOR KO increase the expression of endogenous L1s. RT-qPCR experiments were performed as in (Fig. 1f), but using clonal KO K562 clones instead of mutant cell pools. $n = 2$ biological replicates \times 3 technical replicates (center value as median). The primers do not target the L1-GFP reporter and the cell lines were not dox-induced, so these RT-qPCR assays will not detect L1-GFP transcripts. h. Western blots showing depletion effects of MORC2, MPP8 and TASOR in the mutant pools of K562 cells (left) and in the mutant pools of H9 hESCs without transgenic L1 reporters (right). Two experiments repeated independently with similar results. i. Northern blots showing increased transcription from the L1-GFP reporter in KO K562 clones (same cell lines as in panel c) after 5 days' dox-induction. Two experiments repeated independently with similar results. As observed in Fig. 2b, while HUSH KO significantly increases L1-GFP transcription, MORC2 KO leads to only a modest increase. This is probably because the L1-GFP reporter does not contain the native L1 5' UTR sequence, where MORC2 intensively binds (See Extended Data Fig. 7f,g). The 5 kb and 1.9 kb marks on the membrane refer to the 28S rRNA and 18S rRNA bands respectively. j. Northern blots showing that disruption of MORC2, MPP8 and TASOR increases the expression level of endogenous L1Hs in hESCs, same cell lines as in panel h). Size marker indicated as in panel i). Two experiments repeated independently with similar results. k. Western blots showing protein abundance of L1_ORF1p and HSP90 in the mutant pools of K562 cells and hESCs (same cell line as shown in panel h). Two experiments repeated independently with similar results. Experiments were performed without dox-induction of the transgenic L1 reporter. Due to the strong signal of bands from the KO samples, the blots were exposed for a very short time and the band signal in the Ctrl samples were relatively very weak compared to the KO samples; same case for panels i, j).



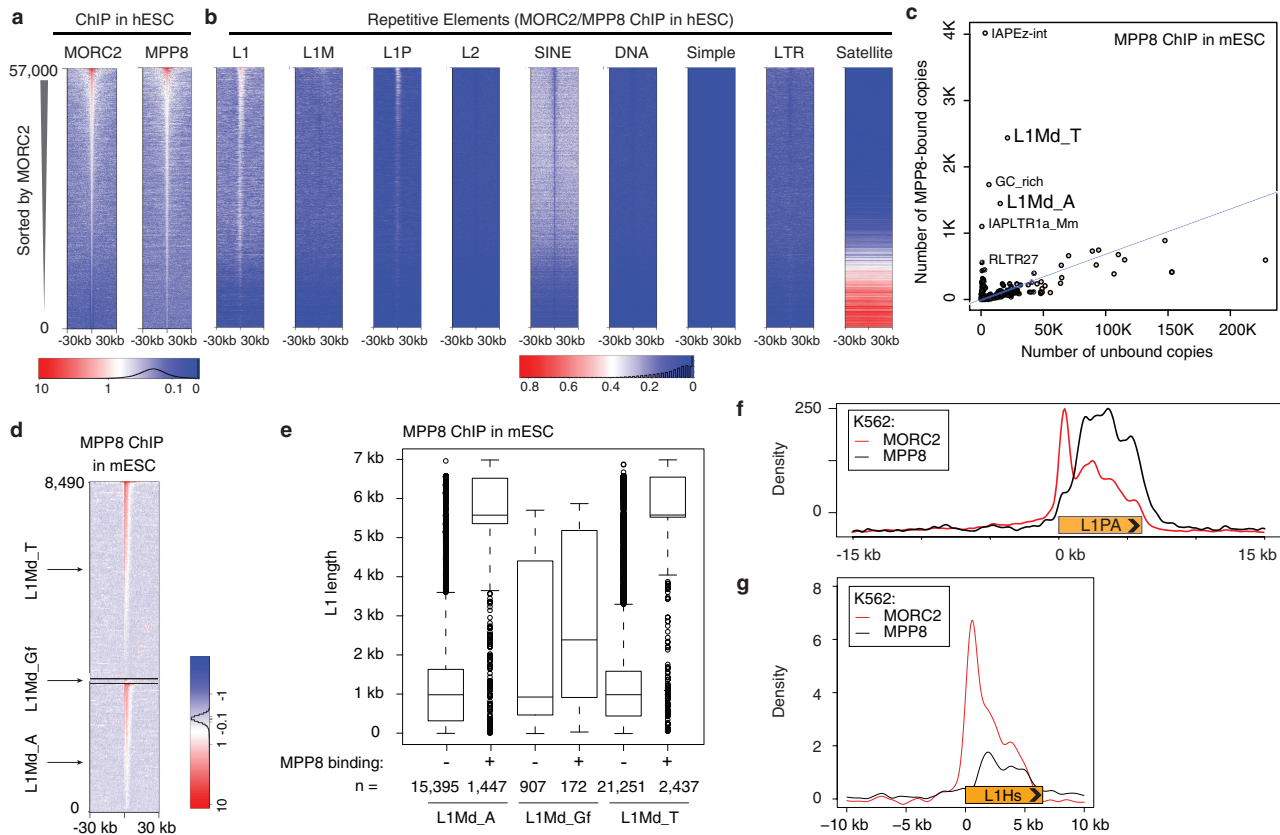
Extended Data Figure 5 | The binding profiles of MORC2, MPP8 and TASOR revealed by ChIP-seq in K562 cells. **a.** Using a paired-end sequencing strategy for the ChIP-seq, together with the sequence divergence within native L1 elements, we could map ChIP-seq reads to individual L1 instances in the genome. Genome browser snapshots of MORC2 ChIP-seq reads alignment over L1PA7 (left) and L1Hs (right). Experiment was repeated once with similar results. Color scale indicates mapping quality score (MAPQ) for each read pair. $\text{MAPQ} = 10 \log_{10} p$, where p is the probability that true alignment belongs elsewhere. With the exception of L1Hs, which is the youngest and least sequence divergent family, the bodies of L1 repeats are uniquely mappable. In case of L1Hs, the 5'UTR is still mappable to determine the level of L1Hs in Ctrl and KO clones. **b.** Genome browser snapshots for MPP8 (blue), TASOR (orange) and MORC2 (purple) ChIP-seq read densities from Ctrl and corresponding KO K562 clones at two representative example genomic loci. Experiment was repeated once with similar results. LINE element occurrences are indicated by blue rectangles at the bottom of the plot. Four instances of long L1 elements are named indicating L1 families they

belong to. Note complete absence of ChIP-seq signal from KO lines and selectivity toward some but not other L1 instances. Of note, while MPP8 and MORC2 ChIP signals were robust, TASOR ChIPs showed relatively weak enrichments (either due to poor antibody quality or genuine biological properties); for this reason, a subset of our downstream analyses is focused on MORC2 and MPP8. **c.** In addition to full length L1, HUSH complex and MORC2 bind 3'UTRs of KRAB Zinc Finger (ZNF) genes. Genome browser snapshots of ChIP-seq read densities over representative examples, from both Ctrl and corresponding KO K562 clones. Experiment was repeated once with similar results. **d.** HUSH complex and MORC2 preferentially bind expressed KRAB-ZNF genes over other ZNF genes. Heatmaps of MPP8 (left) and MORC2 (center) signals over 2,600 ZNF genes, centered in the 3' end of the genes and sorted first by the presence of KRAB domain and then by MPP8 ChIP signal. Upper 1,600 genes are KRAB-ZNF, lower 1,000 non-KRAB ZNF genes. Right heatmaps codes absolute expression level of each gene in RPKM scale from the K562 RNA-seq data (rightmost panel).



Extended Data Figure 6 | HUSH and MORC2 collaborate at binding target L1s. a. Representative genome browser view of normalized ChIP-seq read densities over L1 elements. Experiment was repeated once with similar results. Loss of MPP8 and TASOR results in no detectable binding by MORC2, MPP8 and TASOR, while loss of MORC2 results in partially diminished recruitment of HUSH complex subunits. b. Heatmaps of MPP8 (left), TASOR (center) and MORC2 (right) ChIP-seq signals subtracted

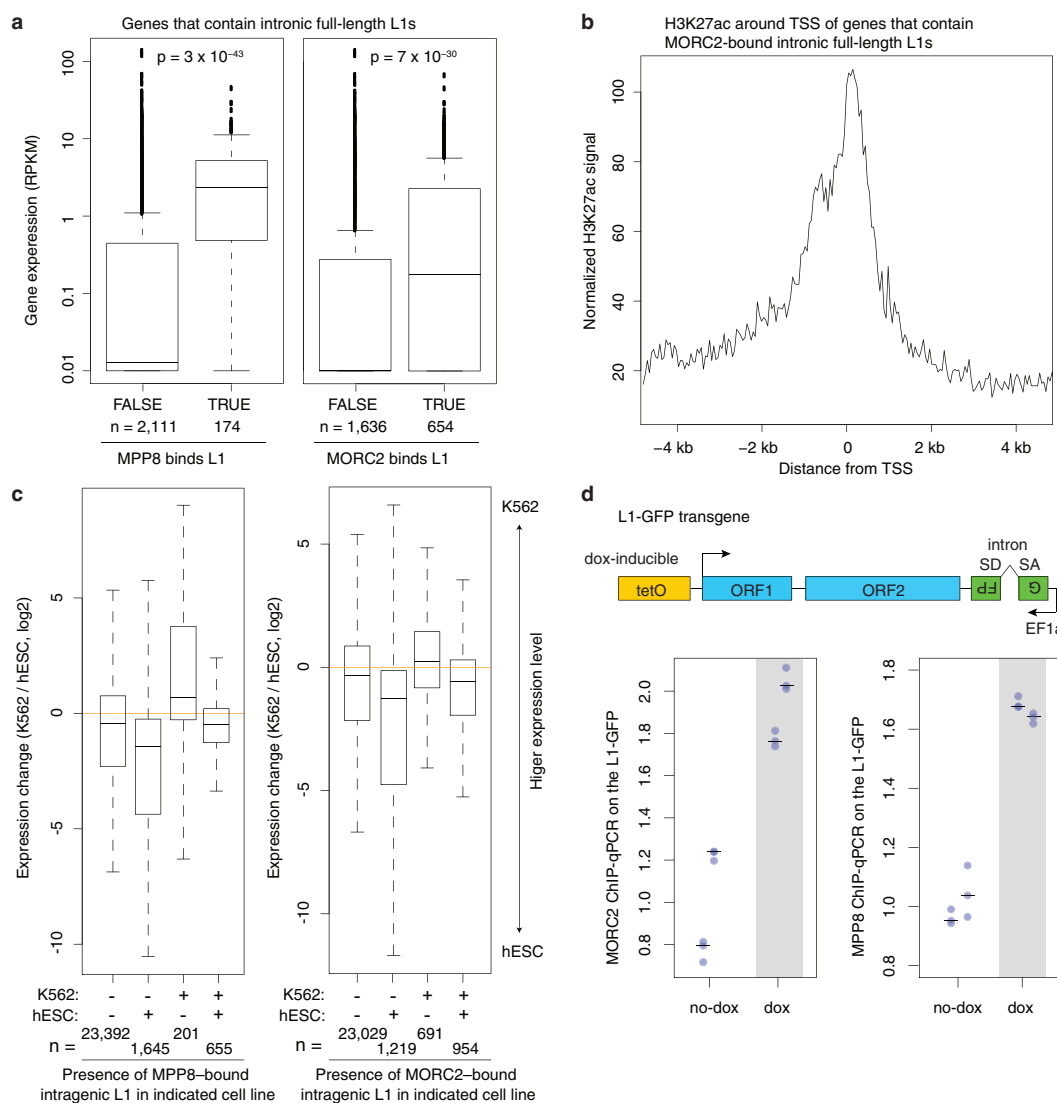
for ChIP signal from corresponding KO lines. Heatmaps are centered on MPP8 and MORC2 peaks, separated by the presence or absence of underlying L1 and then sorted by MPP8 ChIP signal strength. Loss of MORC2 has only partial effect on recruitment of MPP8 and TASOR to the L1 elements, while loss of either MPP8 or TASOR abrogates MORC2 recruitment.



Extended Data Figure 7 | HUSH/MORC2 preferentially bind full-length L1 instances in human ESCs, mouse ESCs and K562 cells.

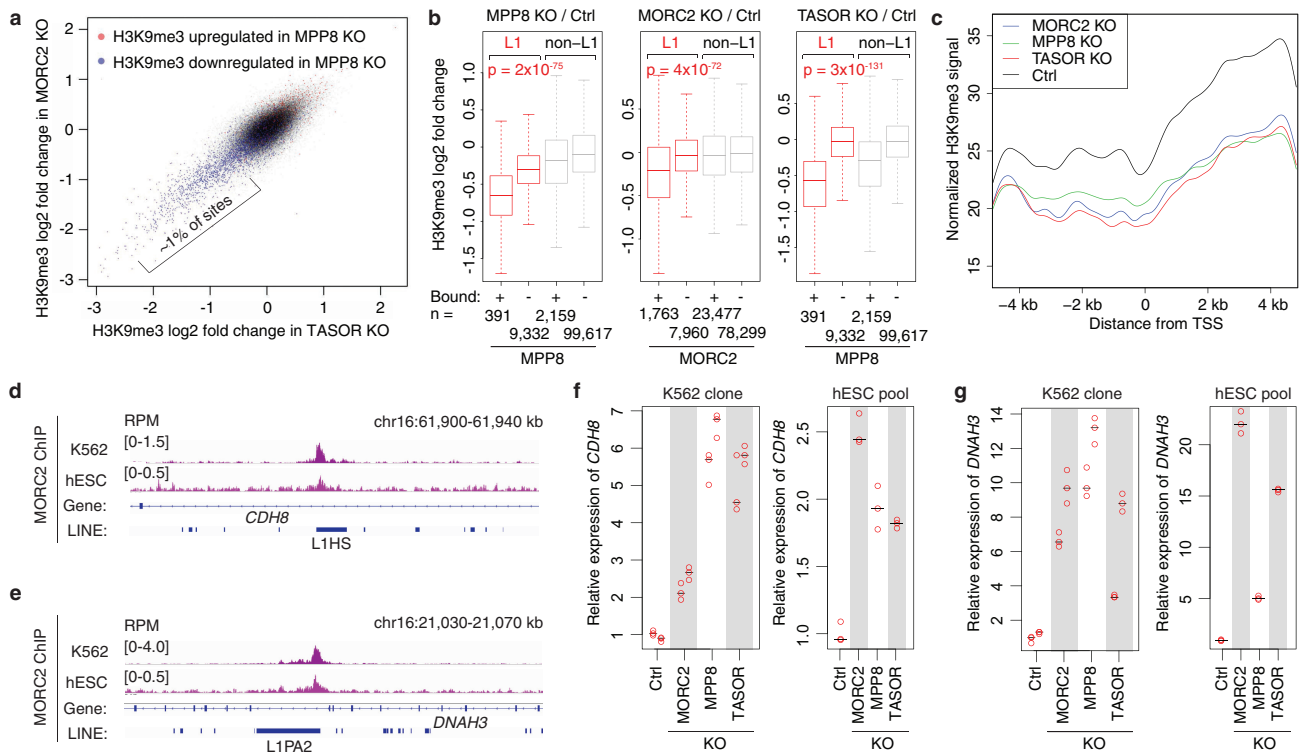
a. Widespread genomic co-binding of MPP8 and MORC2 in hESCs. Heatmap representation of ChIP-seq results at 57,000 genomic loci, centered on MPP8 and MORC2 summits and sorted by MORC2 ChIP-seq signal. Plotted is normalized ChIP read density from hESCs. **b.** Heatmaps of MORC2/MPP8 ChIP-seq density over indicated repeat classes, centered and sorted as in panel **a**. HUSH complex and MORC2 bind predominantly to L1 elements in hESCs, in particular to the primate-specific L1P families, suggesting that HUSH/MORC2-dependent silencing is relevant in many embryonic and somatic cell types. **c.** L1 families that encompass active L1 copies, such as L1Md-T and L1Md-A, are significantly enriched among MPP8 binding sites in mouse ESC. L1Md-Gf is also enriched but not shown due to the low number of instances. Thus, HUSH-mediated L1 regulation appears to be conserved among species. Of note, MPP8 is also strongly enriched at IAP elements, a class of murine endogenous

retroviruses that remain currently mobile in the mouse genome. **d.** MPP8 ChIP-seq heatmaps in mESCs featuring retrotransposition-competent L1Md-T, L1Md-A and L1Md-Gf. **e.** MPP8 preferentially bind full-length L1Md-A and L1Md-T in mESCs. Plotted is size distribution of the indicated L1 instances that overlap with MPP8 ChIP-seq peaks, or remaining L1s that do not overlap with such ChIP-seq signals. Box plots show median and interquartile range (IQR), whiskers are $1.5 \times$ IQR. **f.** Aggregate plots of MORC2 (red) and MPP8 (black) ChIP-seq signals over 500 full-length, MPP8-bound L1PAs, centered on the L1 5' end. **g.** Aggregate plots of MORC2 (red) and MPP8 (black) ChIP-seq signals on L1Hs (L1PA1). Similar as the binding profile on L1PA (panel **f**), MPP8/MORC2 occupy the whole body of L1Hs, with MORC2 additionally binding L1Hs 5'UTR. Please note that ChIP-seq fragments are much less likely to be uniquely mapped, and thus removed by the alignment criteria, within the L1Hs non-5'UTR region, due to their minimal sequence divergence (Extended Data Fig. 5a).



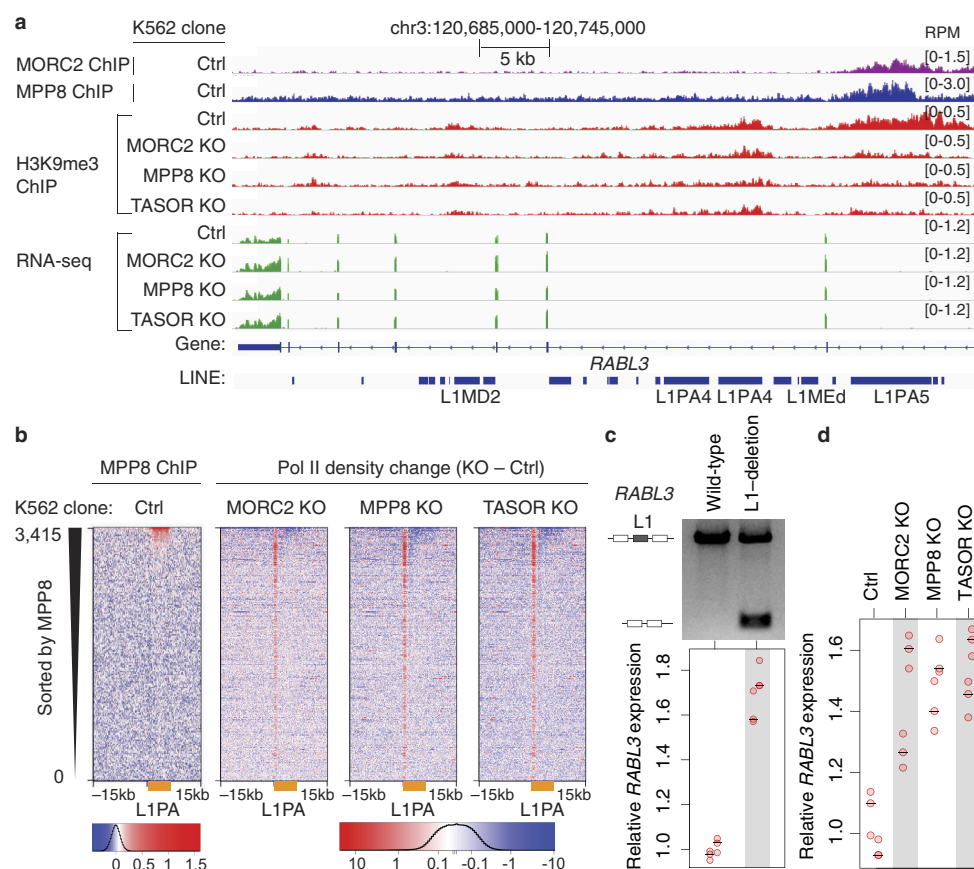
Extended Data Figure 8 | HUSH/MORC2 preferentially bind intronic L1s within actively transcribed genes. a. Genes that contain MPP8 or MORC2 bound intronic L1s are expressed at significantly higher levels in Ctrl K562 cells, compared to genes that contain intronic full-length L1s unbound by MPP8 or MORC2. p-value, two-sided Mann-Whitney-Wilcoxon test. Box plots show median and interquartile range (IQR), whiskers are $1.5 \times$ IQR. b. The promoters of genes that contain MPP8 or MORC2 bound intronic full-length L1s are marked by transcriptionally permissive H3K27ac in wild-type K562 cells. H3K27ac ChIP-seq data are taken from K562 epigenome pilot study, accession number PRJEB8620.

TSS, transcription start site. c. Genes selectively occupied by MORC2/MPP8 either in K562 or in hESC cells exhibit higher gene expression in the corresponding cell line (p-values = 4.3×10^{-107} for MPP8 binding; p-values = 5.0×10^{-92} for MORC2 binding, Kruskal-Wallis test). Boxplots defined as in panel a. RNA-seq datasets for hESC are from SRA entries SRR2043329 and SRR2043330. d. ChIP-qPCR assays quantifying HUSH/MORC2 binding to an inducible L1 transgene in K562 cells before or after its transcriptional induction via Dox. Transcriptional induction increases binding of MORC2 and MPP8 to the L1 transgene. $n = 2$ biological replicates $\times 3$ technical replicates (center value as median).



Extended Data Figure 9 | HUSH/MORC2 facilitate H3K9me3 at their L1 targets for transcription repression. **a.** Concordant subset (~1%) of ($n = 111,499$) H3K9me3 sites in the genome lose H3K9me3 signal in MORC2 KO, MPP8 KO and TASOR KO K562 clones. Two independent lines each for WT, MORC2KO, TASOR KO, MPP8 KO. Plotted is log₂ fold change in H3K9me3 ChIP signal in TASOR KO relative to Ctrl (x-axis) and log₂ fold change in H3K9me3 ChIP signal in MORC2 KO relative to Ctrl (y-axis). Points are color coded with blue sites having significant H3K9me3 loss in MPP8 KO, red sites significantly gaining the signal in MPP8 KO, while gray have no detectable change. Sites that significantly lose H3K9me3 signal in KO line are more likely to have corresponding loss in other KO lines. Odds ratios: 26.23 with 95% confidence intervals (CI) [23.66, 29.10] for MORC2 versus MPP8; 21.70 with 95% CI [19.75, 23.83] for TASOR versus MPP8; 122.53 with 95% CI [109.21, 137.43] for TASOR versus MORC2. $P = 0$ each case, two-sided Fisher's exact test. **b.** Genomic sites that exhibit the strongest loss of H3K9me3 in MORC2, MPP8 or TASOR KOs are preferentially L1 occupied by these factors. Boxplots of log₂ fold change in H3K9me3 relative to Ctrl for MPP8 KO (left), MORC2 KO (center) and TASOR KO (right). Box plots show median and interquartile range (IQR), whiskers are $1.5 \times$ IQR. MPP8 and MORC2

bound L1s show significant loss of H3K9me3 (p-values, two-sided Mann-Whitney-Wilcoxon test). **c.** Averaged distribution of H3K9me3 ChIP-seq signals in Ctrl and KO K562 clones over the host genes that contain the MORC2-targeted intronic full-length L1s, centered on the transcription start site (TSS) of the host genes. **d.** Genome browser showing MORC2 binding at the intronic full-length L1s within *CDH8* in both K562 and hESCs. Experiment was repeated once with similar results. **e.** Genome browser showing MORC2 binding at the intronic full-length L1PA2 within *DNAH3* in both K562 and hESCs. Experiment was repeated once with similar results. **f.** Depletion of MORC2/HUSH increases the expression of *CDH8* in both K562 ($n = 2$ biological replicates \times 3 technical replicates) and hESCs ($n = 3$ technical replicates), as measured by RT-qPCR assay. The *CDH8* expression level was normalized to beta-actin mRNA. All samples were then normalized to Ctrl sample. Center value as median. **g.** Depletion of MORC2/HUSH increases the expression of *DNAH3* in both K562 ($n = 2$ biological replicates \times 3 technical replicates) and hESCs ($n = 3$ technical replicates), as measured by RT-qPCR assay. The *DNAH3* expression level was normalized to beta-actin mRNA. All samples were then normalized to Ctrl sample. Center value as median.



Extended Data Figure 10 | HUSH/MORC2 binding at intronic L1s results in the decreased expression of active host genes. a. Genome browser tracks illustrating loss of HUSH/MORC2 causing decreased H3K9me3 over the intronic L1PA5 element and concomitant increase in the expression of host gene *RABL3*. Experiment was repeated once with similar results. b. Loss of HUSH/MORC2 leads to increased Pol II signals at 5'UTR and decreased Pol II signals within L1 bodies at HUSH-bound L1PA elements (orange bars). Heatmaps show Pol II density change in KO K562 clones compared to Ctrl, centered on the L1 5' end and sorted by MPP8 ChIP signal. c. Deletion of the intronic L1 within *RABL3* causes

increased *RABL3* expression. Upper panel: an agarose gel analysis of the PCR assay with primers flanking the HUSH/MORC2-bound intronic L1; two experiments repeated independently with similar results. Lower panel: RT-qPCR analysis of *RABL3* expression. The *RABL3* expression level was normalized to beta-actin mRNA. All samples were then normalized to wild-type sample. $n = 2$ biological replicates \times 3 technical replicates (center value as median). d. Depletion of MORC2, MPP8, TASOR increases *RABL3* expression. RT-qPCR data normalized as in panel c). $n = 2$ biological replicates \times 3 technical replicates (center value as median).

Opening of the human epithelial calcium channel TRPV6

Luke L. McGoldrick^{1,2*}, Appu K. Singh^{1*}, Kei Saotome¹, Maria V. Yelshanskaya¹, Edward C. Twomey^{1,2}, Robert A. Grassucci¹ & Alexander I. Sobolevsky¹

Calcium-selective transient receptor potential vanilloid subfamily member 6 (TRPV6) channels play a critical role in calcium uptake in epithelial tissues^{1–4}. Altered TRPV6 expression is associated with a variety of human diseases⁵, including cancers⁶. TRPV6 channels are constitutively active^{1,7,8} and their open probability depends on the lipidic composition of the membrane in which they reside; it increases substantially in the presence of phosphatidylinositol 4,5-bisphosphate^{7,9}. Crystal structures of detergent-solubilized rat TRPV6 in the closed state have previously been solved^{10,11}. Corroborating electrophysiological results³, these structures demonstrated that the Ca^{2+} selectivity of TRPV6 arises from a ring of aspartate side chains in the selectivity filter that binds Ca^{2+} tightly. However, how TRPV6 channels open and close their pores for ion permeation has remained unclear. Here we present cryo-electron microscopy structures of human TRPV6 in the open and closed states. The channel selectivity filter adopts similar conformations in both states, consistent with its explicit role in ion permeation. The iris-like channel opening is accompanied by an α -to- π -helical transition in the pore-lining transmembrane helix S6 at an alanine hinge just below the selectivity filter. As a result of this transition, the S6 helices bend and rotate, exposing different residues to the ion channel pore in the open and closed states. This gating mechanism, which defines the constitutive activity of TRPV6, is, to our knowledge, unique among tetrameric ion channels and provides structural insights for understanding their diverse roles in physiology and disease.

We expressed the full-length human TRPV6 (hTRPV6) channel in HEK 293 cells, where it exhibited typical Ca^{2+} permeability¹² (Fig. 1a, b) and current–voltage relationships^{13–17} (Extended Data Fig. 1a). To structurally characterize hTRPV6, we purified it separately in nanodiscs and amphipols (see Methods) and solved the corresponding structures using cryo-electron microscopy (cryo-EM; Extended Data Figs 2, 3 and Extended Data Table 1) to 3.6 Å and 4.0 Å, respectively. Although the reconstructions in nanodiscs and amphipols were nearly identical, the structure solved in nanodiscs had better overall resolution and will be our primary descriptor of hTRPV6. Two-dimensional class averages showed diverse orientations and easily discernible secondary structure features (Fig. 1c). The resulting 3D reconstruction (Fig. 1d, e) showed higher resolution features for the core of the molecule than for its periphery (Extended Data Fig. 2c) and was of sufficient quality (Extended Data Fig. 4) to build each subunit (residues 28–638) of the hTRPV6 homotetramer *de novo*.

The structure of hTRPV6 (Fig. 2a, b) has the same overall architecture as that of rat TRPV6 (rTRPV6)¹⁰. Whereas no discernible lipid densities were observed in the crystal structures of rTRPV6^{10,11}, the hTRPV6 cryo-EM reconstruction reveals 16 (4 per subunit) well-resolved non-protein densities that are intercalated in subunit interfaces and are likely to represent lipids (Fig. 2c). Similarly positioned densities

in the structure of TRPV1¹⁸ were modelled with phosphatidylinositol, phosphatidylcholine and phosphatidylethanolamine lipids. Of the four putative lipid densities in hTRPV6, the fourth density has a clear

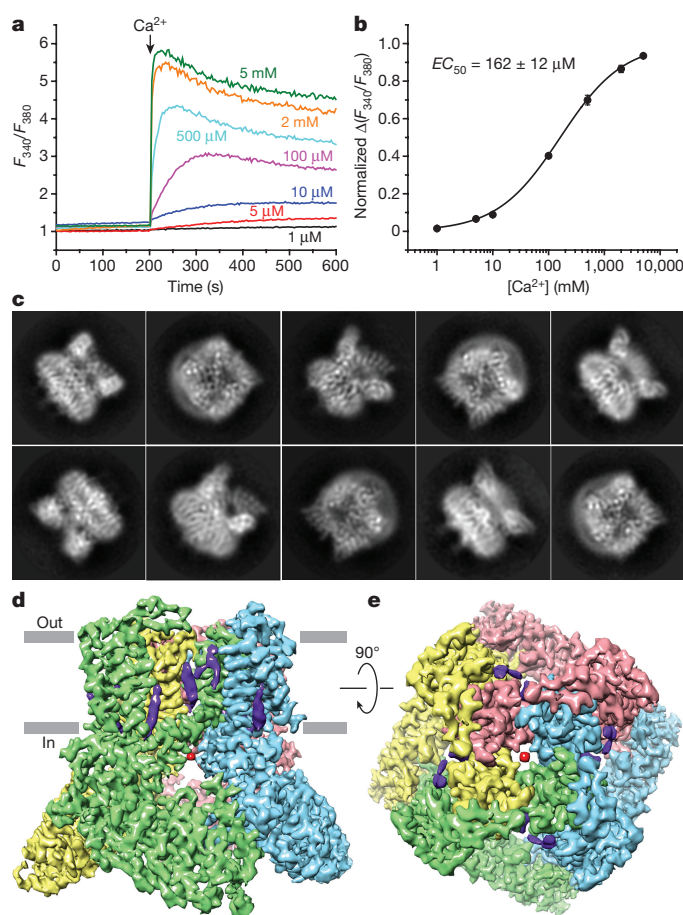


Figure 1 | Function and cryo-EM of hTRPV6. **a, b**, Functional characterization of hTRPV6 using ratiometric fluorescence measurements. **a**, Fluorescence curves recorded from HEK 293 cells expressing hTRPV6 in response to the application of Ca^{2+} (arrow) at different concentrations. These experiments were repeated independently three times with similar results. **b**, Ca^{2+} dose–response curve for the maximal value of fluorescence fitted with the logistic equation. The calculated half maximal effective concentration (EC_{50}) is shown as mean \pm s.e.m. ($n=3$). **c**, Two-dimensional class averages of hTRPV6 particles, showing diverse orientations. **d, e**, hTRPV6 3.6 Å cryo-EM reconstruction, with density shown at 0.035 threshold level (UCSF Chimera) representing hTRPV6 subunits coloured green, cyan, pink and yellow, lipid in purple and ions in red.

¹Department of Biochemistry and Molecular Biophysics, Columbia University, 650 West 168th Street, New York, New York 10032, USA. ²Integrated Program in Cellular, Molecular, and Biomedical Studies, Columbia University, 650 West 168th Street, New York, New York 10032, USA.

*These authors contributed equally to this work.

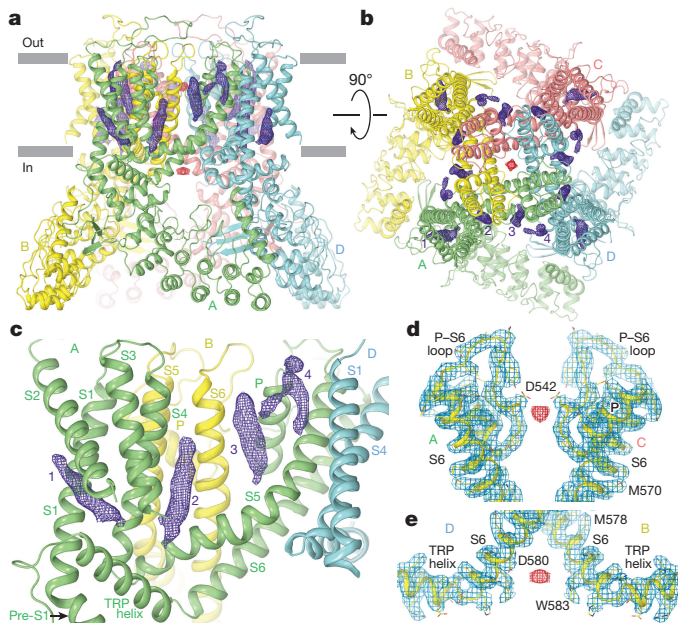


Figure 2 | Structure of hTRPV6. **a, b**, Side (**a**) and top (**b**) views of hTRPV6 tetramer, with each subunit (A–D) shown in a different colour. Putative lipid densities at 3.5σ and ion densities at 4σ are illustrated by purple and red mesh, respectively. **c**, Expanded view of the four (1–4) putative lipid densities per hTRPV6 subunit. **d, e**, Expanded views of the putative ion densities at 4σ at the selectivity filter (**d**) and S6 helices bundle crossing (**e**).

head-and-two-tails appearance. Fitting different lipid molecules into density 4 (Extended Data Fig. 5a–c) suggests that the chemical environment around the lipid head group, including the negatively charged aspartate D525 and polar Y349, Y509, Q513 and Y524 residues, supports binding of phosphatidylethanolamine or phosphatidylcholine rather than phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂). Densities 1–3 have sausage-like appearances and might represent a wider variety of lipid-like molecules, including cholesterol or cholesterol hemisuccinate (CHS) (Extended Data Fig. 5d, e). In physiological conditions, some of these sites can bind PtdIns(4,5)P₂^{7,9}. For example, the positively charged R470 and K484 and polar T479, Q483 and Q596 residues around density 2 create a permissive chemical environment for the negatively charged head group of PtdIns(4,5)P₂. However, the poor fit of PtdIns(4,5)P₂ into density 2 (Extended Data Fig. 5f) suggests that in our cryo-EM structure, density 2 represents a different molecule.

In the crystal structure of rTRPV6 in the closed state, the M577 side chains form a hydrophobic ‘seal’ on the cytoplasmic side of the S6 helices^{10,11}. By contrast, interatomic distances within the pore of the new structure confirmed that the hTRPV6 channel pore is open (Fig. 3). The pore surface is lined by the side chains of D542, T539, N572, I575, D580 and W583, as well as the backbone-carbonyl oxygens of I540, I541 and G579. The narrowest part of the upper pore, the selectivity filter, is formed by the D542 side chains, one from each subunit, which project towards the centre of the pore (Fig. 3a, c, f). We propose that D542 in hTRPV6, similar to D541 in rTRPV6, plays an important role in Ca²⁺ permeation by directly coordinating dehydrated Ca²⁺ ions¹¹. The narrowest part of the hTRPV6 lower pore (9.6 Å interatomic distance) is defined by the side chains of I575 at the S6 bundle crossing (Fig. 3a, c, d). This part of the pore is comparable in size to the pore of open TRPV1 (9.3 Å interatomic distance, measured between side chains of I679)¹⁸ and is wide open for conductance of hydrated Na⁺ or Ca²⁺ ions.

Along the axis of the hTRPV6 pore, there is a strong density about 3.9 Å away from the side chains of D542 (Fig. 2d) that is likely to represent a Ca²⁺ ion bound to a site homologous to the main

Ca²⁺-binding site (site 1) at D541 in the pore of rTRPV6^{10,11}. An additional strong density along the hTRPV6 pore is observed at the bundle crossing of the S6 helices, about 8.0 Å away from D580 and about 6.6 Å away from W583, suggesting that these residues may play an important role in ion permeation (Fig. 2e). Indeed, W583 is conserved in TRPV6 and TRPV5 channels and is involved in the regulation of calcium uptake, as shown by mutation W583A in TRPV5, which induces cell death due to increased calcium influx¹⁹. The density at the S6 helices bundle crossing, which was not observed in the pore of the closed-state rTRPV6^{10,11}, is likely to represent another permeant ion bound in the open pore of hTRPV6.

To determine the structure of hTRPV6 in the closed state, we decided to shift the open–closed state gating equilibrium towards the closed state by interfering with channel activation. Because the open probability of TRPV6 is strongly dependent on membrane lipids^{7,9}, altering lipid binding by mutagenesis might result in channel closure. TRPV1, for example, is activated through an intramembrane vanilloid-binding site (Fig. 4a), which accommodates agonists, such as resiniferatoxin (RTX) and capsaicin, and antagonists, such as capsazepine (CPZ)¹⁸. In the absence of ligands, this site is occupied by the lipid phosphatidylinositol, which favours the closed pore conformation¹⁸. The TRPV1 vanilloid-binding site coincides with hTRPV6 lipid density 2, which may represent the binding site for natural lipid agonists (Fig. 4b).

To test whether this site is critical for channel activation, we mutated R470 to glutamate (R470E). An analogous mutation has previously been shown to eliminate capsaicin-evoked currents in TRPV1²⁰. The mutant hTRPV6(R470E) channels were functional (Extended Data Fig. 1b) but their calcium uptake was about ten times slower than that of wild-type channels (Extended Data Fig. 1e, f), consistent with their less frequent openings. In addition, 2-APB, a TRPV6 inhibitor that acts through the membrane, showed increased affinity to and decreased maximum inhibition of hTRPV6(R470E) compared to wild-type channels (Extended Data Fig. 1i–k), consistent with the R470E mutation altering regulation of TRPV6 by lipids.

We solved the hTRPV6(R470E) structure in amphipols by cryo-EM to 4.2 Å resolution (Extended Data Fig. 6). Consistent with the idea that the site 2 density represents an activating lipid, this density was smaller in hTRPV6(R470E) (Fig. 4c) than in hTRPV6 (Fig. 4b). Confirming that the physical occupancy of site 2 differed, the side chain of Q483 in hTRPV6(R470E) has an altered conformation that would cause clashing with the lipid density in wild-type hTRPV6 (Fig. 4b, c). Supporting the role of Q483 in lipid recognition, an hTRPV6(Q483A) mutant, while being functional (Extended Data Fig. 1c), showed an approximately five times slower calcium uptake than wild-type channels (Extended Data Fig. 1e, g). Notably, the ion channel in hTRPV6(R470E) appears to be closed (Fig. 3b, c, e). Indeed, the size of the pore at the S6 bundle crossing becomes comparable to the narrowest point of the selectivity filter. While the latter is formed by the side chains of D542, which directly coordinate calcium ions for selective permeation, the S6 bundle crossing is formed by the side chains of L574 and M578, which create a hydrophobic seal impermeable to ions and water, and therefore represent the channel gate.

The closed-state structure of hTRPV6(R470E) is nearly identical to the closed-state crystal structure of rat TRPV6 (rTRPV6)¹⁰ and their superposition yields a root mean square deviation (r.m.s.d.) of 0.917 Å. To verify that the rTRPV6 crystal structure represents the physiologically relevant conformation, we solved the structure of rTRPV6 by cryo-EM to 3.9 Å using a lipid nanodisc preparation similar to that used for hTRPV6 (Extended Data Fig. 7). Strikingly, the cryo-EM structure of rTRPV6 is nearly identical (r.m.s.d. = 0.781 Å) to the crystal structure of rTRPV6 (Extended Data Fig. 8a–c). As the hTRPV6(R470E) structure is nearly identical to both the rTRPV6 cryo-EM structure (Extended Data Fig. 8d, e, r.m.s.d. = 0.932 Å) and the rTRPV6 crystal structure, we contend that it represents the true closed state of hTRPV6. Consistently, a much weaker density at site 2 in the cryo-EM structure of rTRPV6 (Fig. 4d) suggests either lower occupancy or greater

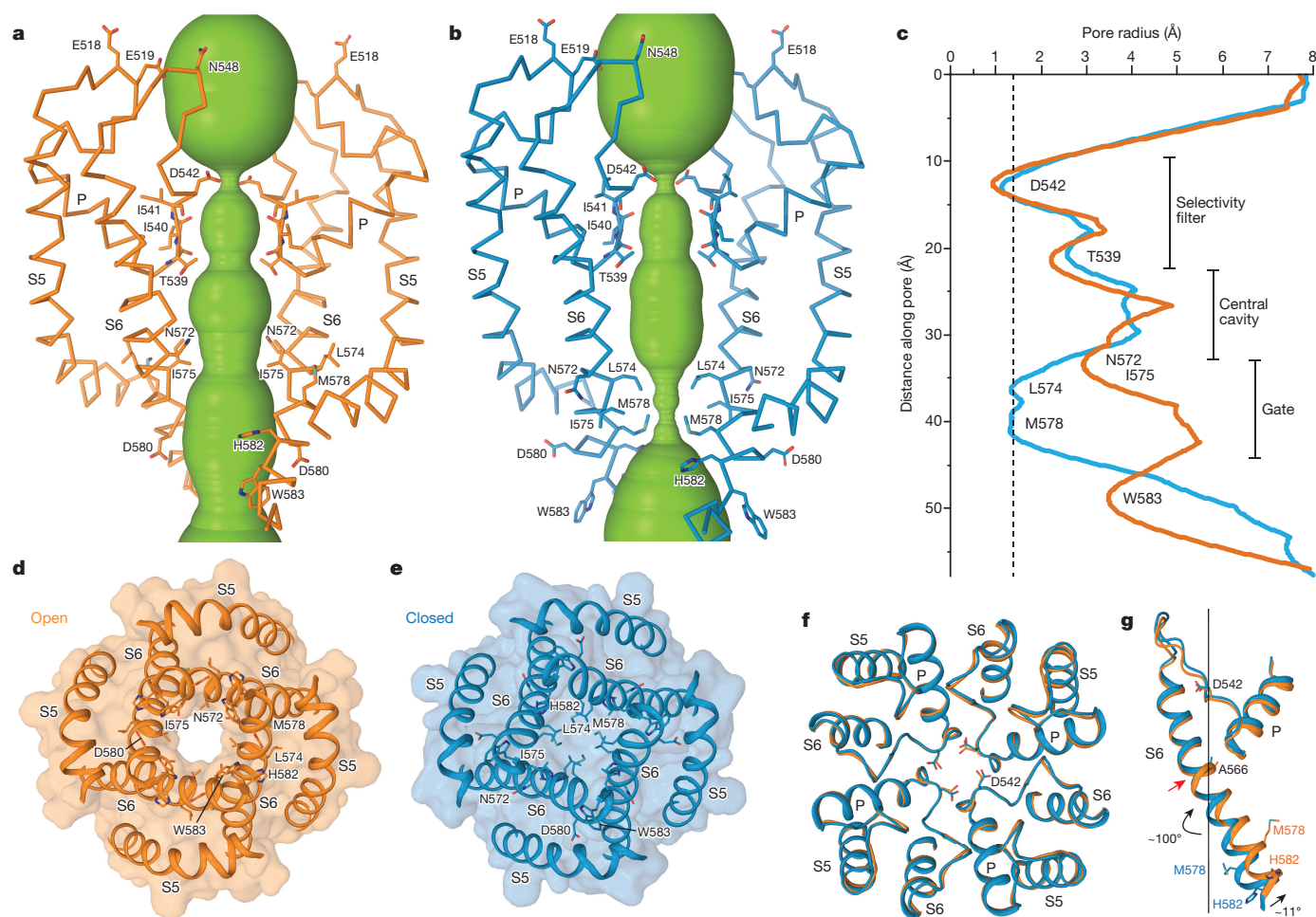


Figure 3 | Open and closed ion channel pore. **a, b**, Ion conduction pathway (green) in open hTRPV6 (**a**) and closed hTRPV6(R470E) (**b**), with residues lining the selectivity filter and around the gate shown as sticks. Only two of four subunits are shown, with the front and back subunits removed for clarity. **c**, Pore radius calculated using HOLE²⁹ for hTRPV6 (orange) and hTRPV6(R470E) (blue). Dashed line corresponds to 1.4 Å (radius of a water molecule). **d, e**, Intracellular view of the S6 bundle

crossing in hTRPV6 (**d**) and hTRPV6(R470E) (**e**). **f**, Superposition of the selectivity filter regions in hTRPV6 (orange) and hTRPV6(R470E) (blue), viewed extracellularly. **g**, Superposition of the P loop and S6 in hTRPV6 (orange) and hTRPV6(R470E) (blue), viewed parallel to the membrane. The straight line shows the pore axis, red arrow indicates the position of the gating hinge alanine A566 and black arrows illustrate ~100° rotation and ~11° bending away from the pore axis of the C-terminal part of S6.

mobility of the putative bound lipid. Because rTRPV6 and hTRPV6 were purified in similar conditions, have 89% overall sequence identity, and have identical amino acid compositions of their site 2 lipid-binding pockets, it remains unclear why one channel was closed and the other open. For example, some lipids within the membranes of the

protein-expressing HEK 293 cells may be important for opening of hTRPV6 but not rTRPV6. Different conformations of rTRPV6 and hTRPV6 might also reflect the ease with which these constitutively active channels rapidly transition between gating states, and that very subtle changes can push this equilibrium towards one state or the other.

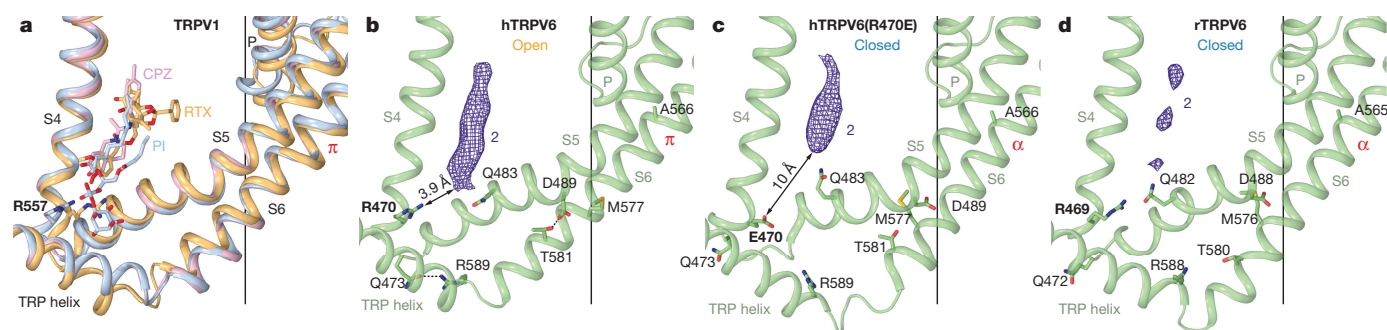


Figure 4 | Activation-related lipid binding pocket. **a**, Superposition of the agonist binding site in TRPV1 structures in the phosphatidylinositol (PI)-bound closed state (blue, PDB ID: 5IRZ), antagonist CPZ-bound closed state (pink, PDB ID: 5IS0) and agonist RTX-bound open state (orange, PDB ID: 5IRX). **b–d**, Putative activating lipid binding site in

open hTRPV6 (**b**), closed hTRPV6(R470E) (**c**) and closed rTRPV6 (**d**), with densities filtered to the same resolution (4.24 Å) and shown at 5.3 σ as purple mesh. Residues involved in gating are shown as sticks. Dashed lines in **b** indicate bonds between the residues.

Such subtle changes, for instance, can originate from different interactions of the membrane-mimicking environment (amphipols or nanodiscs) with helices S1–S3. These helices contain the largest number of membrane lipid-facing residues (69%), of which only 80% are identical between rTRPV6 and hTRPV6.

To understand the structural changes that occur during TRPV6 opening, we compared our hTRPV6(R470E) and hTRPV6 structures. The principal changes occur in the pore-lining helix S6 and originate at A566, which is highly conserved in TRPV5 and TRPV6 (Extended Data Fig. 9k) and located right below the selectivity filter (Fig. 3g). Confirming its important role in gating, substitution of A566 with threonine, the homologous residue conserved in TRPV1–4, greatly reduced the TRPV6 current amplitude (Extended Data Fig. 1d) and slowed calcium uptake by approximately 30 times (Extended Data Fig. 1e, h). Upon opening, S6, which has an α -helical conformation in the closed state, undergoes a local transition to a π -helix. Notably, such a transition has been hypothesized previously, based on a comparison between the TRPV1 and TRPV2 structures²¹. Concurrently, the lower part of S6, which forms the gate in the closed state, rotates by about 100° and bends away from the pore axis by about 11° (Fig. 3g). These rearrangements not only widen the pore for permeant ions but also change the set of residues that face the pore axis (for example, N572 and I575 in the open state compared to L574 and M578 in the closed state). Alanine A566, therefore, acts as a hinge to allow TRPV6 gating at the S6 bundle crossing without changing the conformation of the selectivity filter (Fig. 3f). Correspondingly, the selectivity filter appears to play a crucial role in TRPV6 channel ion permeation rather than gating. Gating-related conformational changes induced by the α -to- π -helical transition in S6 seem to involve only the intracellular portions of S5 and S6, the S4–S5 linker and the TRP helix. Indeed, superposition of the corresponding regions (residues 469–500 plus 566–608) in hTRPV6 and hTRPV6(R470E) gives an r.m.s.d. of 1.74 Å, while the rest of the molecules superpose with a much lower r.m.s.d. of 0.218 Å.

Within the regions involved in gating, pore opening in hTRPV6 is accompanied by the formation of two electrostatic bonds per subunit (Fig. 4b). A salt bridge forms between Q473 in the S4–S5 elbow and R589 in the TRP helix, and a hydrogen bond forms between D489 in the S5 helix and T581 in the S6 helix. Neither interaction is present in the closed-state structures of hTRPV6(R470E) or rTRPV6 and the formation of the hydrogen bond (D489–T581 in hTRPV6(R470E) or D488–T580 in rTRPV6) is prevented by the side chain of M577 or M576, respectively (Fig. 4c, d, Extended Data Fig. 8f, g). The importance of the D489–T581 interaction for hTRPV6 opening is supported by the previous observation that a mutation equivalent to T581A reduces the excessive constitutive activity of an hTRPV6(G516S) mutant²². We speculate that formation of the electrostatic bonds compensates for the energetic cost of the unfavourable α -to- π -helical transition in S6 during channel opening. This structural solution therefore maintains the relative stabilities and similar energy levels of both gating states and supports the constitutive activity of TRPV6. Accordingly, the open and closed conformations of TRPV6 remain in a readily tunable equilibrium that can be shifted towards either state by different stimuli, including lipids^{7,9}.

Our structures of hTRPV6 reveal a gating mechanism that is novel among tetrameric ion channels (Fig. 5, Supplementary Video 1). Although other representatives of the TRP channel family have a local α -to- π -helical transition in the middle of S6^{18,23–25}, they lack the alanine gating hinge (Extended Data Fig. 9). As a result, S6 maintains its secondary structure throughout the entire TRPV1 gating cycle, the same residues face the pore in the closed and open states, and pore widening is observed at both the S6 bundle crossing and the selectivity filter¹⁸. On the other hand, K⁺ channels do have a gating hinge in their pore-forming inner helices^{26,27}. However, this hinge is formed by a glycine located one residue C-terminally compared to the gating hinge alanine in TRPV6 and permits bending of the inner helices by about 30° without an α -to- π transition. The glycine hinge, like the

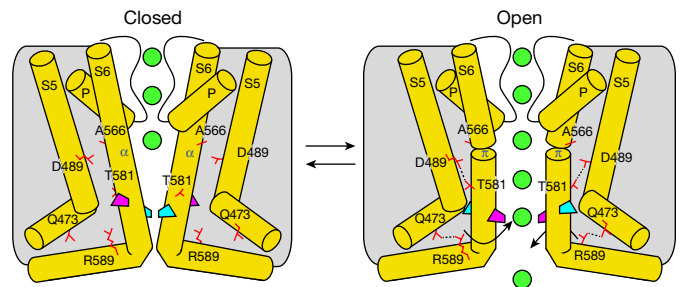


Figure 5 | TRPV6 channel gating mechanism. Cartoons represent the structural changes associated with TRPV6 channel gating. Transition from the closed to open state, stabilized by the formation of salt bridges (dashed lines), leads to permeation of ions (green spheres) and is accompanied by a local α -to- π -helical transition in S6 that maintains the selectivity filter conformation, while the lower part of S6 bends by about 11° and rotates by about 100°. These movements result in a different set of residues (cyan versus pink symbols) lining the pore in the vicinity of the channel gate.

alanine hinge in TRPV6, allows gating of K⁺ channels to occur at the inner helices bundle crossing without changing the selectivity filter. However, unlike TRPV6, the glycine gating hinge in K⁺ channels does not introduce a 100° rotation of the lower parts of the pore-forming helices and correspondingly does not change the residues that line the pore gate region. An alanine gating hinge is present in the pore-forming helices of ionotropic glutamate receptor (iGluR) family tetrameric ion channels²⁸. However, this alanine gating hinge is located at the ion channel gate region. Correspondingly, bending the pore-forming helices at the iGluR alanine gating hinge directly alters the diameter of the pore in close proximity to the gate without an α -to- π transition. The alanine gating hinge in TRPV5 and TRPV6 channels is therefore a unique structural element that is likely to be associated with their exclusive physiological role as constitutively active calcium uptake channels.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 August; accepted 28 November 2017.

Published online 20 December 2017.

- Peng, J. B. *et al.* Molecular cloning and characterization of a channel-like transporter mediating intestinal calcium absorption. *J. Biol. Chem.* **274**, 22739–22746 (1999).
- Yue, L., Peng, J. B., Hediger, M. A. & Clapham, D. E. CaT1 manifests the pore properties of the calcium-release-activated calcium channel. *Nature* **410**, 705–709 (2001).
- Owsianik, G., Talavera, K., Voets, T. & Nilius, B. Permeation and selectivity of TRP channels. *Annu. Rev. Physiol.* **68**, 685–717 (2006).
- Woudenberg-Vrenken, T. E. *et al.* Functional TRPV6 channels are crucial for transepithelial Ca²⁺ absorption. *Am. J. Physiol. Gastrointest. Liver Physiol.* **303**, G879–G885 (2012).
- Fecher-Trost, C., Weissgerber, P. & Wissenbach, U. TRPV6 channels. *Handb. Exp. Pharmacol.* **222**, 359–384 (2014).
- Lehen'kyi, V., Raphaël, M. & Prevorskaya, N. The role of the TRPV6 channel in cancer. *J. Physiol. (Lond.)* **590**, 1369–1376 (2012).
- Zakharian, E., Cao, C. & Rohacs, T. Intracellular ATP supports TRPV6 activity via lipid kinases and the generation of PtdIns(4,5)P₂. *FASEB J.* **25**, 3915–3928 (2011).
- den Dekker, E., Hoenderop, J. G., Nilius, B. & Bindels, R. J. The epithelial calcium channels, TRPV5 & TRPV6: from identification towards regulation. *Cell Calcium* **33**, 497–507 (2003).
- Velisetty, P. *et al.* A molecular determinant of phosphoinositide affinity in mammalian TRPV channels. *Sci. Rep.* **6**, 27652 (2016).
- Singh, A. K., Saotome, K. & Sobolevsky, A. I. Swapping of transmembrane domains in the epithelial calcium channel TRPV6. *Sci. Rep.* **7**, 10669 (2017).
- Saotome, K., Singh, A. K., Yelshanskaya, M. V. & Sobolevsky, A. I. Crystal structure of the epithelial calcium channel TRPV6. *Nature* **534**, 506–511 (2016).
- Bödding, M. & Flockerzi, V. Ca²⁺ dependence of the Ca²⁺-selective TRPV6 channel. *J. Biol. Chem.* **279**, 36546–36552 (2004).
- Hoenderop, J. G., Nilius, B. & Bindels, R. J. Calcium absorption across epithelia. *Physiol. Rev.* **85**, 373–422 (2005).

14. Hoenderop, J. G. *et al.* Function and expression of the epithelial Ca^{2+} channel family: comparison of mammalian ECaC1 and 2. *J. Physiol. (Lond.)* **537**, 747–761 (2001).
15. Kovacs, G. *et al.* Inhibition of the human epithelial calcium channel TRPV6 by 2-aminoethoxydiphenyl borate (2-APB). *Cell Calcium* **52**, 468–480 (2012).
16. Fecher-Trost, C. *et al.* The *in vivo* TRPV6 protein starts at a non-AUG triplet, decoded as methionine, upstream of canonical initiation at AUG. *J. Biol. Chem.* **288**, 16629–16644 (2013).
17. Voets, T., Janssens, A., Prenen, J., Droogmans, G. & Nilius, B. Mg^{2+} -dependent gating and strong inward rectification of the cation channel TRPV6. *J. Gen. Physiol.* **121**, 245–260 (2003).
18. Gao, Y., Cao, E., Julius, D. & Cheng, Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature* **534**, 347–351 (2016).
19. van der Wijst, J. *et al.* A gate hinge controls the epithelial calcium channel TRPV5. *Sci. Rep.* **7**, 45489 (2017).
20. Boukalova, S., Marsakova, L., Teisinger, J. & Vlachova, V. Conserved residues within the putative S4–S5 region serve distinct functions among thermosensitive vanilloid transient receptor potential (TRPV) channels. *J. Biol. Chem.* **285**, 41455–41462 (2010).
21. Zubcevic, L. *et al.* Cryo-electron microscopy structure of the TRPV2 ion channel. *Nat. Struct. Mol. Biol.* **23**, 180–186 (2016).
22. Hofmann, L., Wang, H., Beck, A., Wissenbach, U. & Flockerzi, V. A conserved gating element in TRPV6 channels. *Cell Calcium* **63**, 24–28 (2017).
23. Paulsen, C. E., Armache, J. P., Gao, Y., Cheng, Y. & Julius, D. Structure of the TRPA1 ion channel suggests regulatory mechanisms. *Nature* **520**, 511–517 (2015).
24. Huynh, K. W. *et al.* Structure of the full-length TRPV2 channel by cryo-EM. *Nat. Commun.* **7**, 11130 (2016).
25. Shen, P. S. *et al.* The structure of the polycystic kidney disease channel PKD2 in lipid nanodiscs. *Cell* **167**, 763–773.e11 (2016).
26. Jiang, Y. *et al.* The open pore conformation of potassium channels. *Nature* **417**, 523–526 (2002).
27. Jiang, Y. *et al.* X-ray structure of a voltage-dependent K^{+} channel. *Nature* **423**, 33–41 (2003).
28. Twomey, E. C., Yelshanskaya, M. V., Grassucci, R. A., Frank, J. & Sobolevsky, A. I. Channel opening and gating mechanism in AMPA-subtype glutamate receptors. *Nature* **549**, 60–65 (2017).
29. Smart, O. S., Neduvellil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.* **14**, 354–360, 376 (1996).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank T. Rohacs for advice on electrophysiological recordings, J. Frank for comments on the manuscript, H. Kao for computational support and members of the E.C. Greene laboratory for assistance with their fluorimeter. L.L.M. and E.C.T. are supported by the NIH (T32 GM008224 and F31 NS093838, respectively). A.I.S. is supported by the NIH (R01 CA206573, R01 NS083660), and Amgen Young Investigator and Irma T. Hirschl Career Scientist awards. Data were collected at the Columbia University Medical Center cryo-EM facility and at the Simons Electron Microscopy Center and National Resource for Automated Molecular Microscopy located at the New York Structural Biology Center, supported by grants from the Simons Foundation (349247), NYSTAR, and the NIH (GM103310).

Author Contributions L.L.M., A.K.S. and A.I.S. designed the project, built models and analysed data. L.L.M. and A.K.S. carried out Fura-2 experiments, cryo-EM data collection and processing. L.L.M., A.K.S., K.S. and M.V.Y. developed expression and purification protocols. L.L.M., A.K.S. and K.S. designed constructs and prepared protein samples. M.V.Y. carried out electrophysiology experiments. E.C.T. advised on cryo-EM workflow. R.A.G. and E.C.T. assisted with microscope operation. A.I.S. supervised the project. L.L.M., A.K.S., K.S., M.V.Y., E.C.T. and A.I.S. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to A.I.S. (as4005@cumc.columbia.edu).

Reviewer Information Nature thanks R. Gaudet, X. Li and Y. Mori for their contribution to the peer review of this work.

METHODS

Construct. The full-length human TRPV6 (residues 1–725) and rat TRPV6 (rTRPV6, residues 1–727) were each introduced into a pEG BacMam vector³⁰, with the C-terminal thrombin cleavage site (LVPRG) followed by the streptavidin affinity tag (WSHPQFEK). The R470E mutation in hTRPV6 was introduced by conventional mutagenesis. The rTRPV6 construct previously used for crystallographic studies, rTRPV6* (ref. 10), was also introduced into a pEG BacMam vector but with eGFP inserted between the thrombin site and the streptavidin tag. Compared to wild-type rTRPV6, the rTRPV6* construct is C-terminally truncated by 59 residues and contains three point mutations in the ankyrin repeat domain (I62Y, L92N and M96Q).

Expression and purification. All constructs were expressed and purified similarly to TRPV6_{Cys11}. Bacmids and baculoviruses were made using a standard method³⁰. The P2 baculovirus, produced in Sf9 cells (Thermo Fisher Scientific, mycoplasma negative, GIBCO #12659017), was added to HEK 293S cells lacking N-acetylglucosaminyltransferase I (GnT1[−]) and grown in suspension (mycoplasma test negative, ATCC #CRL-3022) in Freestyle 293 medium (GIBCO-Life Technologies #12338-018) supplemented with 2% FBS at 37 °C and 5% CO₂. Eight to twelve hours after transduction, 10 mM sodium butyrate was added to enhance protein expression and the temperature was reduced to 30 °C. At 48–72 h post-transduction, cells were harvested by low-speed centrifugation in a Sorvall Evolution RC Centrifuge (Thermo Scientific) at 5,471g for 15 min, washed in phosphate-buffered saline (PBS) pH 8.0, and pelleted in an Eppendorf Centrifuge 5810 at 3,202g for 10 min. The cell pellet was resuspended and subjected to sonication with a Misonix sonicator (12 × 15 s, power level 8) in a buffer containing 150 mM NaCl, 20 mM Tris-HCl (pH 8.0), 1 mM βME (β-mercaptoethanol) and protease inhibitors (0.8 μM aprotinin, 2 μg/ml leupeptin, 2 μM pepstatin A and 1 mM phenylmethylsulfonyl fluoride); 50 ml was used per 800 ml of HEK 293 cell culture. Subsequently, the lysate was clarified after centrifugation using a Sorvall RC-5C Plus centrifuge at 9,900g for 15 min, and the membranes were collected by ultracentrifugation in a Beckman Coulter ultracentrifuge equipped with a Beckman Coulter Type 45 Ti Rotor at 186,000g for one hour. The membranes were then mechanically homogenized, and solubilized for 1–2 h in 150 mM NaCl, 20 mM Tris-HCl pH 8.0, 1% DDM (n-dodecyl-β-D-maltopyranoside), 0.1% CHS, and 1 mM βME. Insoluble material was removed by ultracentrifugation for 40 min in a Beckman Coulter Type 45 Ti Rotor at 186,000g and the supernatant was added to streptavidin-linked resin and rotated for 10–14 h at 4 °C. Next, the resin was washed with 10 column volumes of wash buffer containing 150 mM NaCl, 20 mM Tris-HCl pH 8.0, 1 mM βME, 0.1% DDM, and 0.01% CHS. The bound protein was eluted in wash buffer to which 2.5 mM D-desthiobiotin was added. All constructs were purified by size exclusion chromatography using a Superose 6 column equilibrated in 150 mM NaCl, 20 mM Tris-HCl pH 8.0, 1 mM βME, 0.1% DDM, and 0.01% CHS. Tris(2-carboxyethyl)phosphine (TCEP; 10 mM) was added to the peak fractions, which were pooled and concentrated for channel reconstitution in nanodiscs or amphipols. The rTRPV6 and rTRPV6* constructs were expressed and purified similarly to the hTRPV6 constructs but without the addition of CHS to any buffer. Additionally, after elution from the streptavidin-linked resin, the rTRPV6* fusion protein was concentrated to ~1.0 mg/ml and subjected to thrombin digestion at a mass ratio of 1:100 (thrombin:protein) for one hour at 22 °C with rocking, before size exclusion chromatography. Prior to reconstitution in nanodiscs or amphipols, the concentration of each construct was adjusted to approximately 1.2 mg/ml.

Reconstitution of TRPV6 protein into nanodiscs and amphipols. Both hTRPV6 and rTRPV6 were incorporated into conventional MSP2N2 lipid nanodiscs as described previously¹⁸. In brief, soybean polar lipid extract (Avanti #541602) was solubilized in buffer containing 20 mM Tris pH 8.0, 150 mM NaCl, 2 mM TCEP, and 15 mM DDM to create a 10-mM stock. Purified sample was mixed with the soybean polar lipid extract stock (~7.6 mg/ml) and MSP2N2 (~5.3 mg/ml) at a molar ratio of approximately 1:3:166 for both hTRPV6 (monomer:MSP2N2:lipid) and rTRPV6 (monomer:MSP2N2:lipid) and rocked at room temperature for one hour. Subsequently, 10 mg of Bio-beads SM2 (Bio-rad) pre-wet in buffer (20 mM Tris pH 8.0, 150 mM NaCl, 1 mM βME) was added to 0.5 ml of mixture and the mixture was rotated at 4 °C. After one hour, an additional 10 mg of Bio-beads SM2 was added and the resulting mixture was rotated at 4 °C for ~20 h. The Bio-beads SM2 were removed by pipetting and TRPV6 reconstituted in nanodiscs was isolated from empty nanodiscs by size exclusion chromatography using a Superose 6 column equilibrated in 150 mM NaCl, 20 mM Tris-HCl pH 8.0, and 1 mM βME.

cNW11 circularized nanodiscs were prepared as described previously³¹. Purified rTRPV6* was incorporated into cNW11 (2.0 mg/ml) circularized nanodiscs using the procedure described above for the MSP2N2 nanodiscs but with a molar ratio of 1:10:267 (rTRPV6* monomer:cNW11:lipid).

For reconstitution in A8-35 amphipols (Anatrace #A835), we adapted the previously described procedure³². hTRPV6 or hTRPV6(R470E) was mixed with amphipols at a 1:3 mass ratio (protein:amphipols) and incubated for three hours with rotation at 4 °C. After three hours, 7–8 mg per 0.5 ml of mixture of Bio-beads SM2 pre-wet in buffer containing 20 mM Tris pH 8.0, 150 mM NaCl, 1 mM βME was added to the protein–amphipols mixture to facilitate the reconstitution of TRPV6 into amphipols. The mixture was rotated for ~20 h at 4 °C and the amphipols-solubilized TRPV6 was purified as described above.

Cryo-EM sample preparation and data collection. Au/Au grids were prepared as described³³. In brief, grids were prepared by first coating C-flat (Protochips) CF-1.2/1.3-2Au 200 mesh holey carbon grids with ~50 nm gold using an Edwards Auto 306 evaporator. Subsequently, an Ar/O₂ plasma treatment (6 min, 50 W, 35.0 s.c.c.m. Ar, 11.5 s.c.c.m. O₂) was used to remove the carbon with a Gatan Solaris (model 950) Advanced Plasma Cleaning System. The grids were again plasma treated (H₂/O₂, 20 s, 10 W, 6.4 s.c.c.m. H₂, 27.5 s.c.c.m. O₂) before sample application in order to make their surfaces hydrophilic. A Vitrobot Mark IV (FEI) was used to plunge-freeze the grids after the application of 3 μl protein solution with 100% humidity at 5 °C, a blot time of 2 or 3 s, blot force set to 3, and a wait time of 20 s. A concentration of 0.5 mg/ml was used for the nanodiscs-solubilized protein and 0.3 mg/ml for the amphipols-solubilized protein.

The hTRPV6 in nanodiscs data were collected on a Tecnai F30 Polara (C_s 2.26 mm) at 300 kV equipped with a Gatan K2 Summit direct electron detection (DED) camera (Gatan) using Legion³⁴. We collected 1,733 micrographs in super-resolution mode with a pixel size of 0.98 Å across a defocus range of −1.5 μm to −3.5 μm. The total dose, ~67 e[−] Å^{−2}, was attained by using a dose rate of ~8.0 e[−] pixel^{−1} s^{−1} across 40 frames for 8 s total exposure time. We collected 1,538 hTRPV6 in amphipols micrographs and 1,301 rTRPV6* micrographs as described above. We collected 2,167 rTRPV6 micrographs as described above but in counting mode with a pixel size of 0.98 Å. The hTRPV6(R470E) data were collected on a C_s-corrected Titan Krios (FEI) equipped with a post-column GIF Quantum energy filter at 300 kV. We collected 3,540 micrographs in counting mode with a pixel size of 1.10 Å across a defocus range of −1.5 μm to −3.5 μm. The total dose, ~67 e[−] Å^{−2}, was attained by using a dose rate of ~8.0 e[−] pixel^{−1} s^{−1} across 50 frames for 10 s total exposure time.

Image processing. Data were collected using the Gatan K2 Summit DED camera (Gatan) in super-resolution mode and binned 2 × 2. Frame alignment was done using MotionCorr³⁵. CTF correction, using CTFFIND4³⁶ for the hTRPV6 in nanodiscs dataset and gCTF³⁷ for all other datasets, was performed on non-dose-weighted micrographs and subsequent data processing was done on dose-weighted micrographs. All other data processing, unless stated otherwise, was performed using Relion 2.0³⁸. For each dataset, 1,000–2,000 particles were manually selected to generate 2D classes for use in auto-picking.

In processing the hTRPV6 in nanodiscs dataset, seven 2D classes were used for automatically picking 509,569 particles from the 1,733 collected micrographs. The particle images were binned to a pixel size of 1.96 Å per pixel and screened by 2D classification to remove aberrantly picked particles. The remaining 508,019 particles were subjected to 3D classification into 10 classes with no symmetry imposed. A density map was generated in Chimera from the crystal structure of rTRPV6 (PDB ID: 5IWK), low-pass filtered to 40 Å, and used as an initial reference. Five classes, comprising 313,369 particle images, exhibited structural features of a quality that warranted further processing. Of the five, one showed structural features of higher detail and comprised 71,582 particle images. The particle images composing this class were extracted without binning (0.98 Å per pixel), refined with C4 symmetry using the same reference (unbinned) as the prior round of classification, low-pass filtered to 40 Å, and post-processed. The resulting map was then used as a reference for the second round of 3D classification in which the particles composing the best five aforementioned classes were extracted with binning (1.98 Å per pixel) and split into ten classes with C4 symmetry imposed. Two new classes, comprising a total of 67,034 particles, exhibited structural features of a quality that warranted further processing. These 67,034 particle images were extracted without binning (0.98 Å per pixel), refined with C4 symmetry using the same reference (unbinned) as the prior round of classification low-pass filtered to 40 Å, and post-processed. The resulting map was used for the final round of 3D classification in which the two best classes from the prior round of 3D classification were without binning, and with C4 symmetry imposed, classified into 10. The four best classes, comprising 46,124 particle images, were refined together and post-processed to generate the final 3.6 Å map. This relatively small number of particles compared to the initial pool of picked up 509,569 particles indicates that the majority of picked up particles represent either artefacts or contaminants or TRPV6 molecules in alternative conformations including different gating states or unnatural conformations produced by the artificial environment of the cryo-EM grid.

Each dataset was processed using a workflow similar to that described above and the reported resolutions were estimated using the Fourier shell correlation (FSC = 0.143) criterion on masking-effect-corrected FSC curves calculated between two independent half maps^{39,40}. The local resolutions were estimated with unfiltered half maps using ResMap⁴¹ and EM density maps were visualized using UCSF Chimera⁴².

The cryo-EM data collected for full-length rTRPV6 yielded a low-resolution (6.4 Å) reconstruction that was sufficient to conclude that it represents a closed-state conformation similar to rTRPV6*. As it is lacking high-resolution detail, it is not described in the main text or Extended Data.

Model building. To build the open- and closed-state models of TRPV6 in COOT⁴³, we used the rTRPV6* crystal structure¹⁰ as a guide. The resulting models were refined against unfiltered half maps in real space with constraints using PHENIX⁴⁴. The refined models were tested for overfitting (Extended Data Figs 2f, 3f, 6f, 7f) by shifting their coordinates by 0.5 Å with shake in PHENIX and building their corresponding densities in Chimera⁴² from the shaken models. FSC was calculated between the densities from the shaken models, the half maps used in PHENIX refinement (work), the second half maps (free) and the unfiltered sum maps, using EMAN2⁴⁵. The local resolutions in the transmembrane regions of our hTRPV6 in nanodiscs and hTRPV6(R470E) maps reached 2.5 Å as estimated by ResMap⁴¹. These high resolutions allowed us to unambiguously define the conformation of S6 in the open and closed states as well as the existence of the π -helix in the extracellular half of S6 in the open state. Structures were visualized and figures were prepared in Pymol⁴⁶.

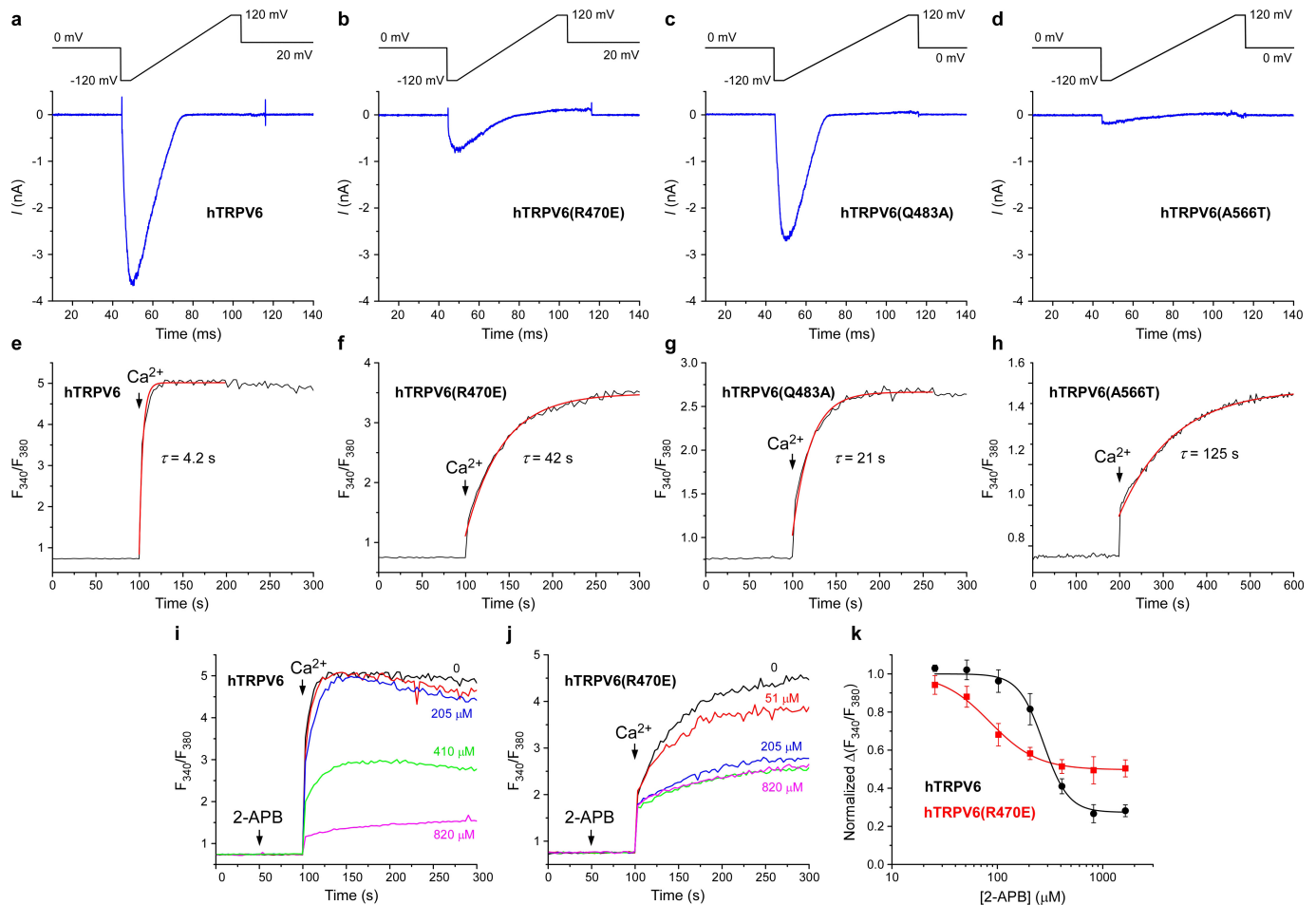
Fura 2-AM measurements. Wild-type hTRPV6 or hTRPV6(R470E) fused to C-terminal streptavidin tag was expressed in HEK 293 cells. Cells were harvested 50–60 h after transduction by centrifugation at 600g for 5 min. The cells were resuspended in pre-warmed modified HEPES-buffered saline (HBS) (118 mM NaCl, 4.8 mM KCl, 1 mM MgCl₂, 5 mM D-glucose, 10 mM HEPES pH 7.4) containing 5 µg/ml Fura2-AM (Life Technologies) and incubated at 37 °C for 45 min. The loaded cells were then centrifuged for 5 min at 600g, resuspended in prewarmed, modified HBS, and incubated again at 37 °C for 25–35 min in the dark. The cells were subsequently pelleted and washed twice, then resuspended in modified HBS for experiments. The cells were kept on ice in the dark for a maximum of ~2 h before fluorescence measurements, which were conducted using a spectrofluorometer QuantaMaster 40 (Photon Technology International) at room temperature in a quartz cuvette under constant stirring. Intracellular Ca²⁺ was measured by taking the ratio of two excitation wavelengths (340 and 380 nm) at one emission wavelength (510 nm). The excitation wavelength was switched at 1-s intervals.

Electrophysiology. HEK 293 cells (ATCC #CRL-1573) were grown on glass cover slips in 35-mm dishes and were transduced with the same P2 virus as was used for large-scale protein production. Recordings were made at room temperature, 36–72 h post-transduction. Currents from whole cells, typically held at a 0 or –60 mV membrane potential, were recorded using an Axopatch 200B amplifier (Molecular Devices, LLC), filtered at 5 kHz and digitized at 10 kHz using a low-noise data acquisition system (Digidata 1440A) and pCLAMP software (Molecular Devices, LLC). The external solution contained (in mM): 140 NaCl, 6 CsCl, 1 MgCl₂, 10 HEPES pH 7.4 and 10 glucose. To evoke monovalent currents, 1 mM EGTA was added to the external solution. The internal solution contained (in mM): 100 CsAsp, 20 CsF, 10 EGTA, 3 MgCl₂, 4 NaATP and 20 HEPES pH 7.2.

TRPV6 currents were recorded in response to 50-ms voltage ramps from –120 mV to 120 mV (see Extended Data Fig. 1). Data analysis was performed using the computer program Origin 9.1.0 (OriginLab Corp.).

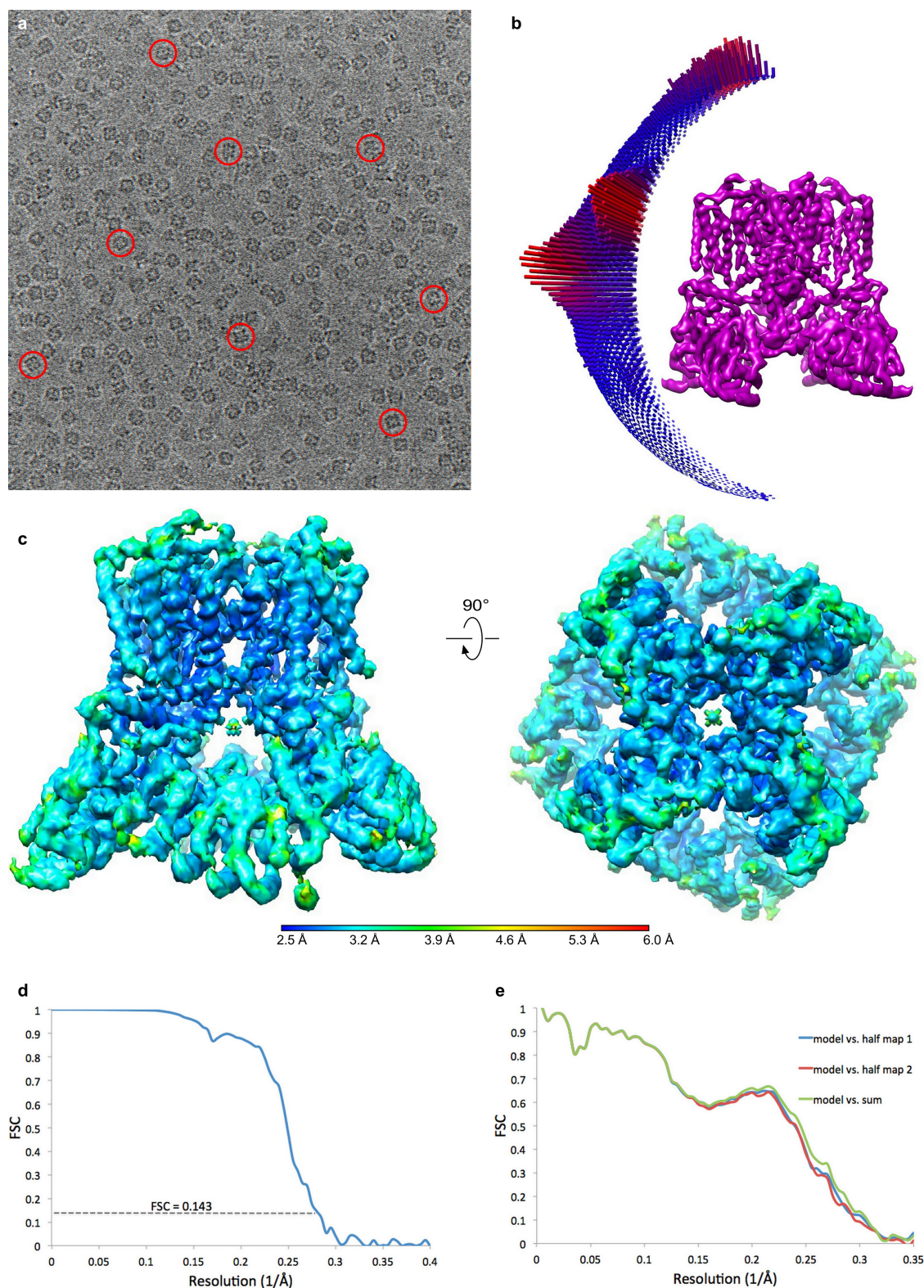
Data availability. Cryo-EM density maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession numbers EMDB-7120 (hTRPV6 in nanodiscs), EMDB-7121 (hTRPV6 in amphipols), EMDB-7122 (hTRPV6(R470E)) and EMDB-7123 (rTRPV6*). Model coordinates have been deposited in the Protein Data Bank (PDB) under accession numbers 6BO8 (hTRPV6 in nanodiscs), 6BO9 (hTRPV6 in amphipols), 6BOA (hTRPV6(R470E)) and 6BOB (rTRPV6*). All other data are available from the corresponding author upon request.

30. Goehring, A. *et al.* Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat. Protocols* **9**, 2574–2585 (2014).
31. Nasr, M. L. *et al.* Covalently circularized nanodiscs for studying membrane proteins and viral entry. *Nat. Methods* **14**, 49–52 (2017).
32. Liao, M., Cao, E., Julius, D. & Cheng, Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* **504**, 107–112 (2013).
33. Russo, C. J. & Passmore, L. A. Ultrastable gold substrates for electron cryomicroscopy. *Science* **346**, 1377–1380 (2014).
34. Suloway, C. *et al.* Automated molecular microscopy: the new Legimon system. *J. Struct. Biol.* **151**, 41–60 (2005).
35. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
36. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
37. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
38. Kimanius, D., Forsberg, B. O., Scheres, S. H. W. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016).
39. Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
40. Chen, S. *et al.* High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35 (2013).
41. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
42. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
43. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
44. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68**, 352–367 (2012).
45. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
46. The PyMOL Molecular Graphics System (DeLano Scientific, 2002).
47. Doyle, D. A. *et al.* The structure of the potassium channel: molecular basis of K⁺ conduction and selectivity. *Science* **280**, 69–77 (1998).
48. Jiang, Y. *et al.* Crystal structure and mechanism of a calcium-gated potassium channel. *Nature* **417**, 515–522 (2002).
49. Long, S. B., Campbell, E. B. & Mackinnon, R. Crystal structure of a mammalian voltage-dependent Shaker family K⁺ channel. *Science* **309**, 897–903 (2005).
50. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).



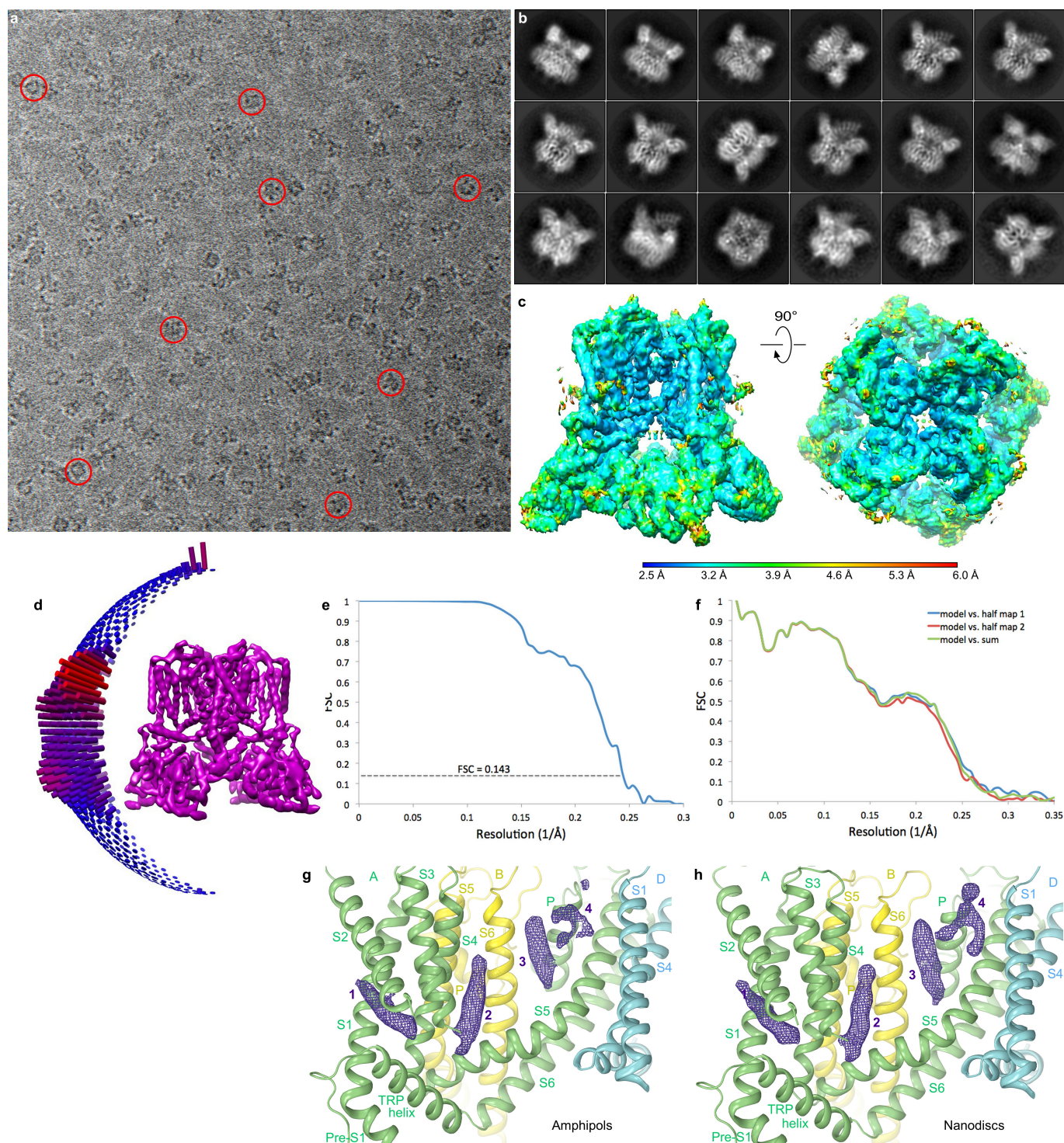
Extended Data Figure 1 | Functional characterization of wild-type and mutant hTRPV6 channels. **a–d**, Whole-cell patch-clamp recordings from HEK 293 cells expressing wild-type hTRPV6 (**a**), hTRPV6(R470E) (**b**), hTRPV6(Q483A) (**c**) and hTRPV6(A566T) (**d**). Leak-subtracted currents (blue) are shown in response to voltage ramp protocols illustrated above the recordings. Although the shapes of the currents for wild-type and mutant hTRPV6 channels were similar, their amplitudes were different. The average current amplitudes at -60 -mV membrane potential (mean \pm s.e.m.) were $3,171 \pm 767$ pA ($n = 11$) for wild-type hTRPV6; 918 ± 267 pA ($n = 9$) for hTRPV6(R470E); $2,239 \pm 398$ pA ($n = 7$) for hTRPV6(Q483A); and 145 ± 52 pA ($n = 5$) for hTRPV6(A566T). **e–h**, Kinetics of calcium uptake using Fura-2 AM ratiometric fluorescence measurements. Representative fluorescence curves are shown for wild-type hTRPV6 (**e**), hTRPV6(R470E) (**f**), hTRPV6(Q483A) (**g**) and hTRPV6(A566T) (**h**) in response to application of 2 mM Ca^{2+} (arrow). Exponential fits are shown in red, with the time constants indicated. Over five measurements, the time constants (mean \pm s.e.m.) were 4.2 ± 0.5 s for hTRPV6; 47 ± 13 s for hTRPV6(R470E); 18.9 ± 0.8 s for hTRPV6(Q483A); and 121 ± 12 s for hTRPV6(A566T). At $n = 5$ and $P = 0.05$, the time constant values for wild-type and mutant channels were statistically different (two-sided t -test). **i, j**, Fluorescence curves for wild-type hTRPV6 (**i**) and hTRPV6(R470E) (**j**) in response to application

of 2 mM Ca^{2+} after pre-incubation of cells in different concentrations of 2-APB. These experiments were repeated independently three times with similar results. **k**, Dose-response curves for 2-APB inhibition calculated for wild-type hTRPV6 (black) and hTRPV6(R470E) (red) ($n = 3$ for all measurements). The changes in the fluorescence intensity ratio at 340 and 380 nm (F_{340}/F_{380}) evoked by addition of 2 mM Ca^{2+} after pre-incubation with various concentrations of 2-APB were normalized to the maximal change in F_{340}/F_{380} after addition of 2 mM Ca^{2+} in the absence of 2-APB. Curves through the data points are fits with the logistic equation, with the mean \pm s.e.m. values of half maximal inhibitory concentration (IC_{50}), 274 ± 27 μM and 85 ± 5 μM , and the maximal inhibition, $72.6 \pm 2.7\%$ and $50.3 \pm 1.1\%$, for hTRPV6 and hTRPV6(R470E), respectively. The leftward shift of the 2-APB dose-response curve of hTRPV6(R470E), when compared to the dose-response curve of wild-type hTRPV6, indicates an increased affinity of the channel for an activating lipid ligand. On the other hand, the reduced maximum inhibition of hTRPV6(R470E) at high concentrations of 2-APB, when compared to that of wild-type hTRPV6, indicates a reduced efficacy of 2-APB that could be a result of the R470E mutation disrupting the mechanism by which 2-APB binding is allosterically coupled to channel gating.



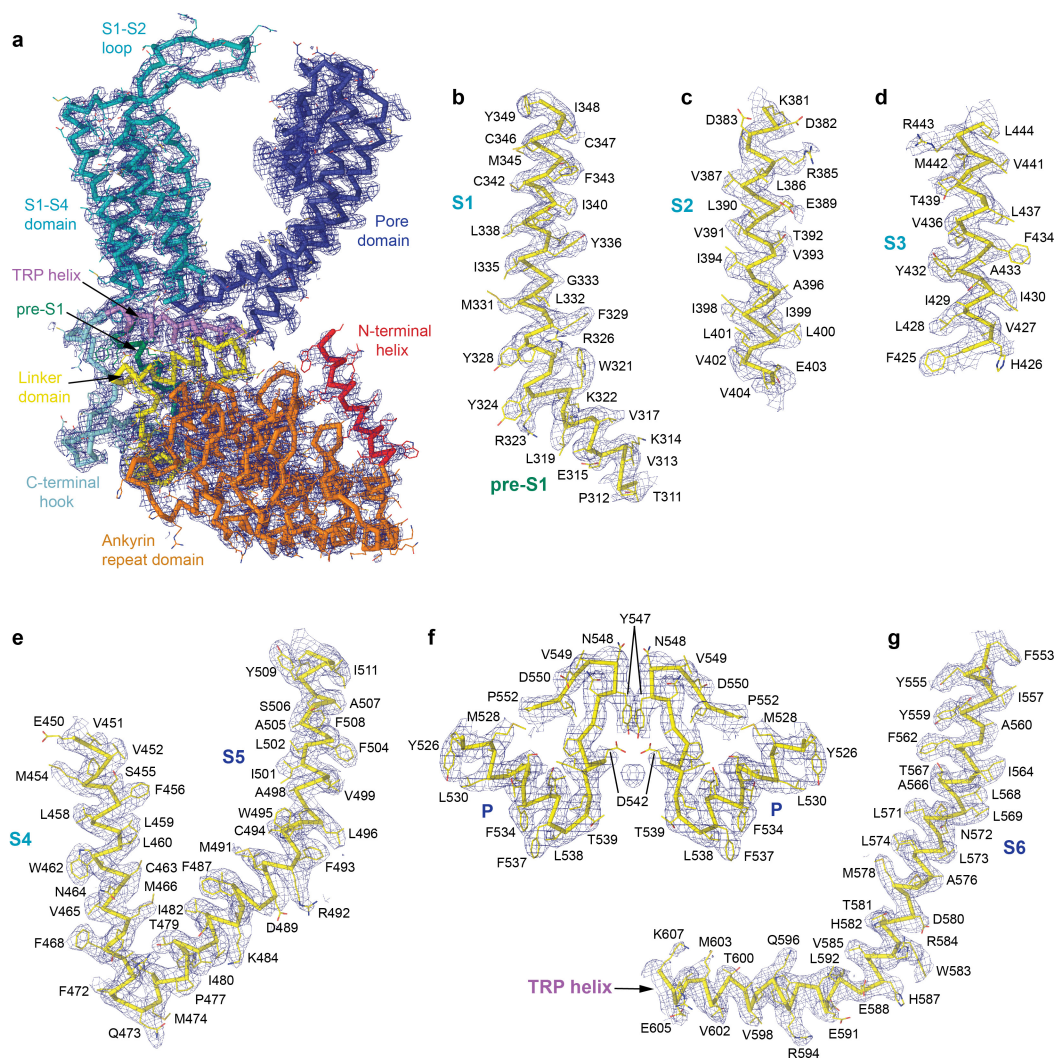
Extended Data Figure 2 | Overview of single-particle cryo-EM for hTRPV6 in nanodiscs. **a**, Example cryo-EM micrograph for hTRPV6 in nanodiscs with example particles circled in red. **b**, Orientations of particles that contributed to the final 3.6 Å reconstruction. Longer red rods represent orientations with more particles. **c**, Local resolution mapped on density at 0.013 threshold level (UCSF Chimera) calculated using Resmap

and two unfiltered half maps, with the highest resolution observed for the channel core. **d**, FSC curve calculated between half maps. **e**, Cross-validation FSC curves for the refined model versus unfiltered half maps (only half map1 was used for PHENIX refinement) and the unfiltered summed map.

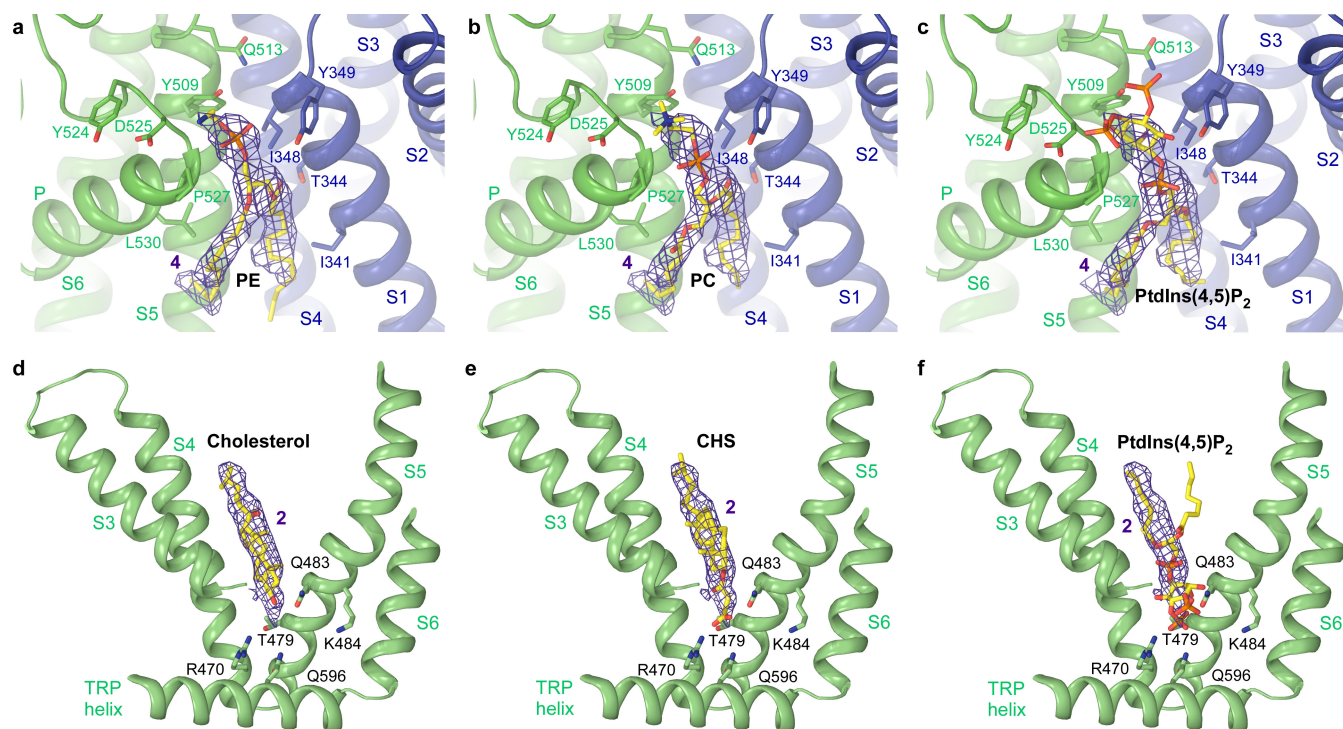


Extended Data Figure 3 | Overview of single-particle cryo-EM for hTRPV6 in amphipols and comparison to the reconstruction in nanodiscs. **a**, Example cryo-EM micrograph for hTRPV6 in amphipols with example particles circled in red. **b**, Reference-free 2D class averages of hTRPV6 in amphipols illustrating different particle orientations. **c**, Local resolution mapped on density at 0.01 threshold level (UCSF Chimera) calculated using Resmap and two unfiltered half maps, with the highest resolution observed for the channel core. **d**, Orientations of particles that

contribute to the final 4.0 Å reconstruction. Longer red rods represent orientations that comprise more particles. **e**, FSC curve calculated between half maps. **f**, Cross-validation FSC curves for the refined model versus unfiltered half maps (only half map 1 was used for PHENIX refinement) and the unfiltered summed map. **g**, **h**, Comparison of putative lipid densities for hTRPV6 in amphipols (**g**) and nanodiscs (**h**), filtered to the same (4.0 Å) resolution and shown at 3.5 σ as purple mesh.



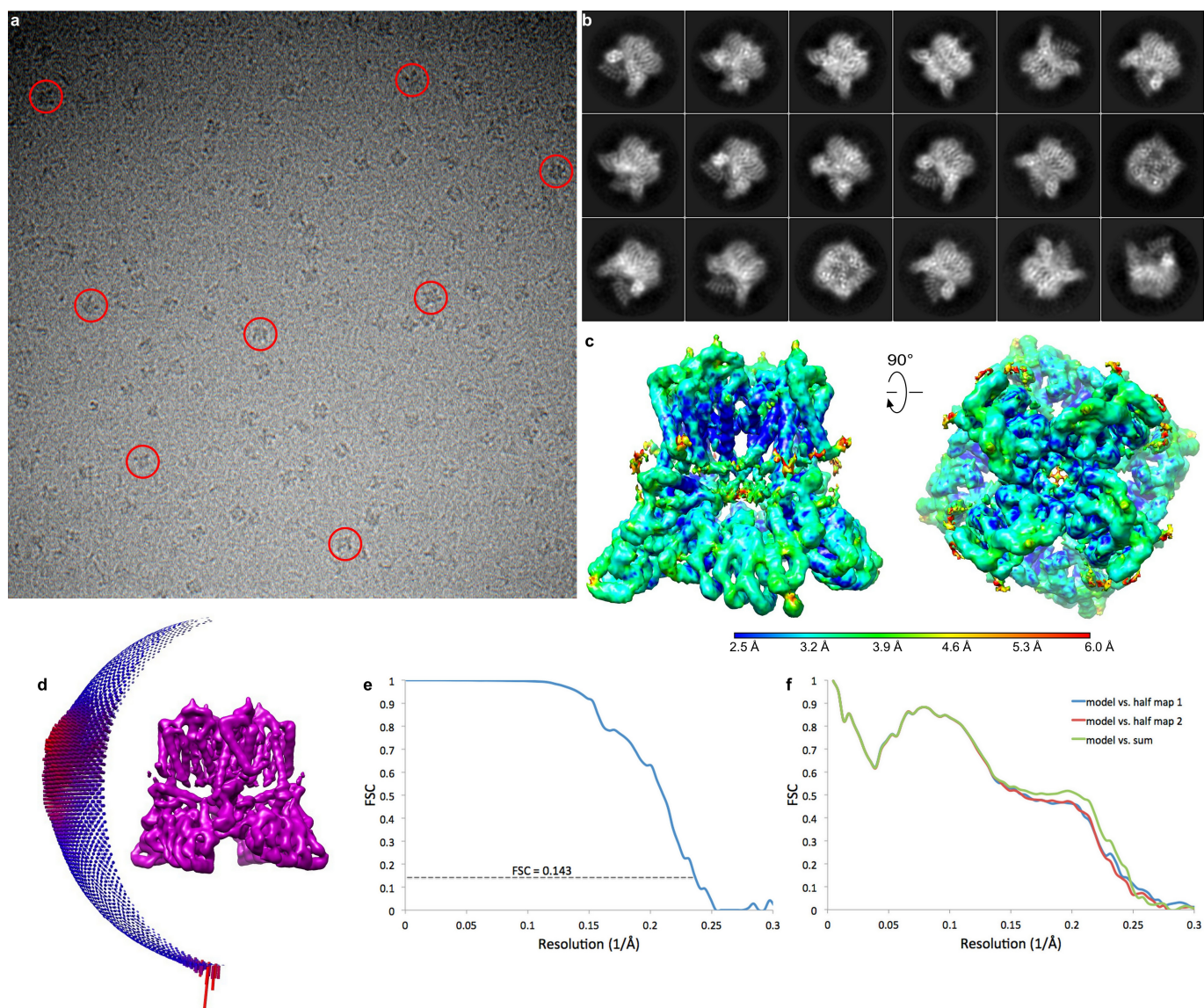
Extended Data Figure 4 | Cryo-EM density for hTRPV6 in nanodiscs. **a**, Cryo-EM density at 4σ for a single hTRPV6 subunit, with the protein shown in ribbon and coloured according to domains. **b–g**, Fragments of the hTRPV6 transmembrane domain with the corresponding cryo-EM densities.



Extended Data Figure 5 | Fitting lipids into cryo-EM density.

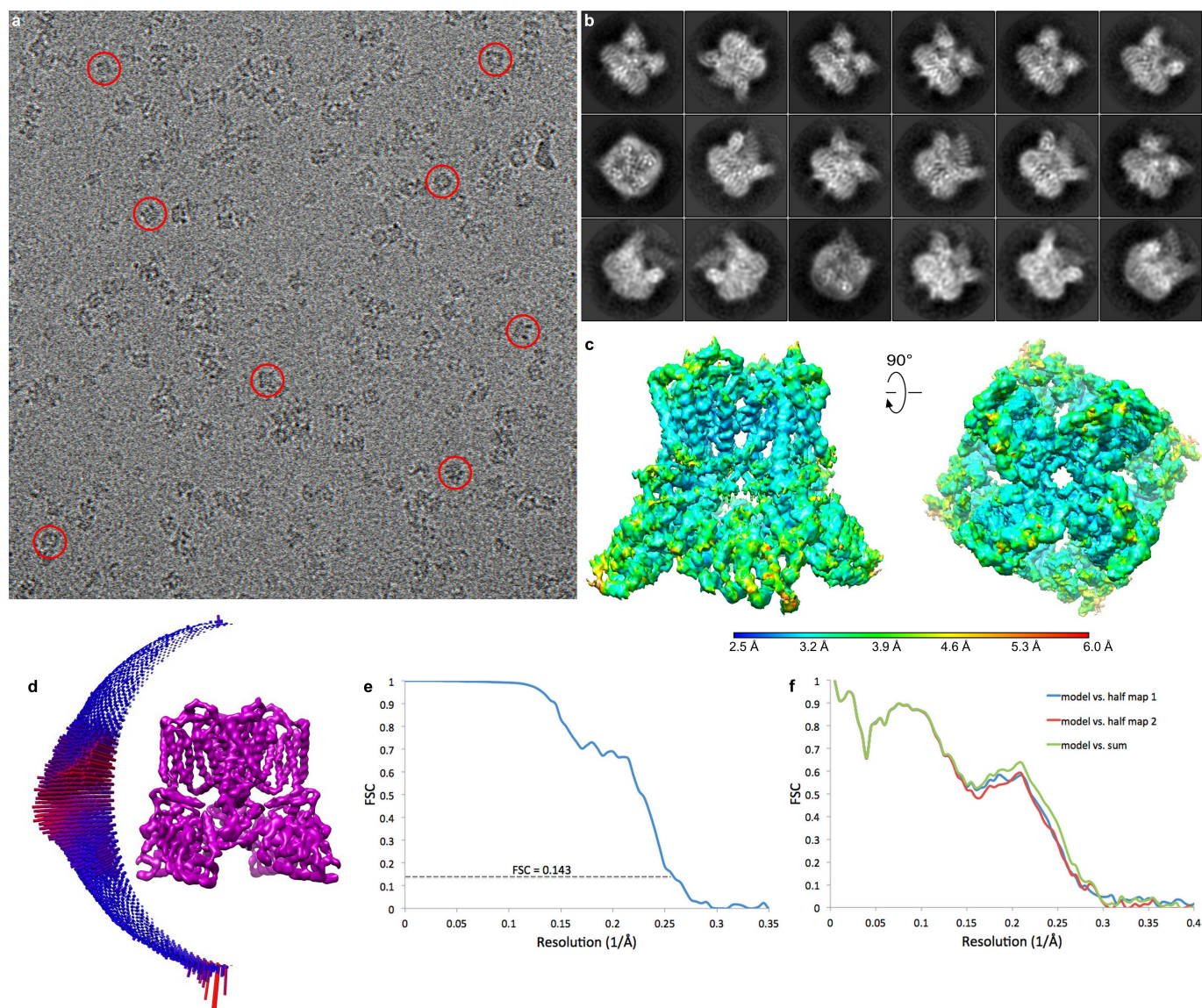
a–c, Molecules of phosphatidylethanolamine (PE, **a**), phosphatidylcholine (PC, **b**) and phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂, **c**) fitted

into the site 4 lipid density shown at 3.5σ as purple mesh. **d–f,** Molecules of cholesterol (**d**), CHS (**e**) and PtdIns(4,5)P₂ (**f**) fitted into the site 2 putative activating lipid density shown at 5.3σ .



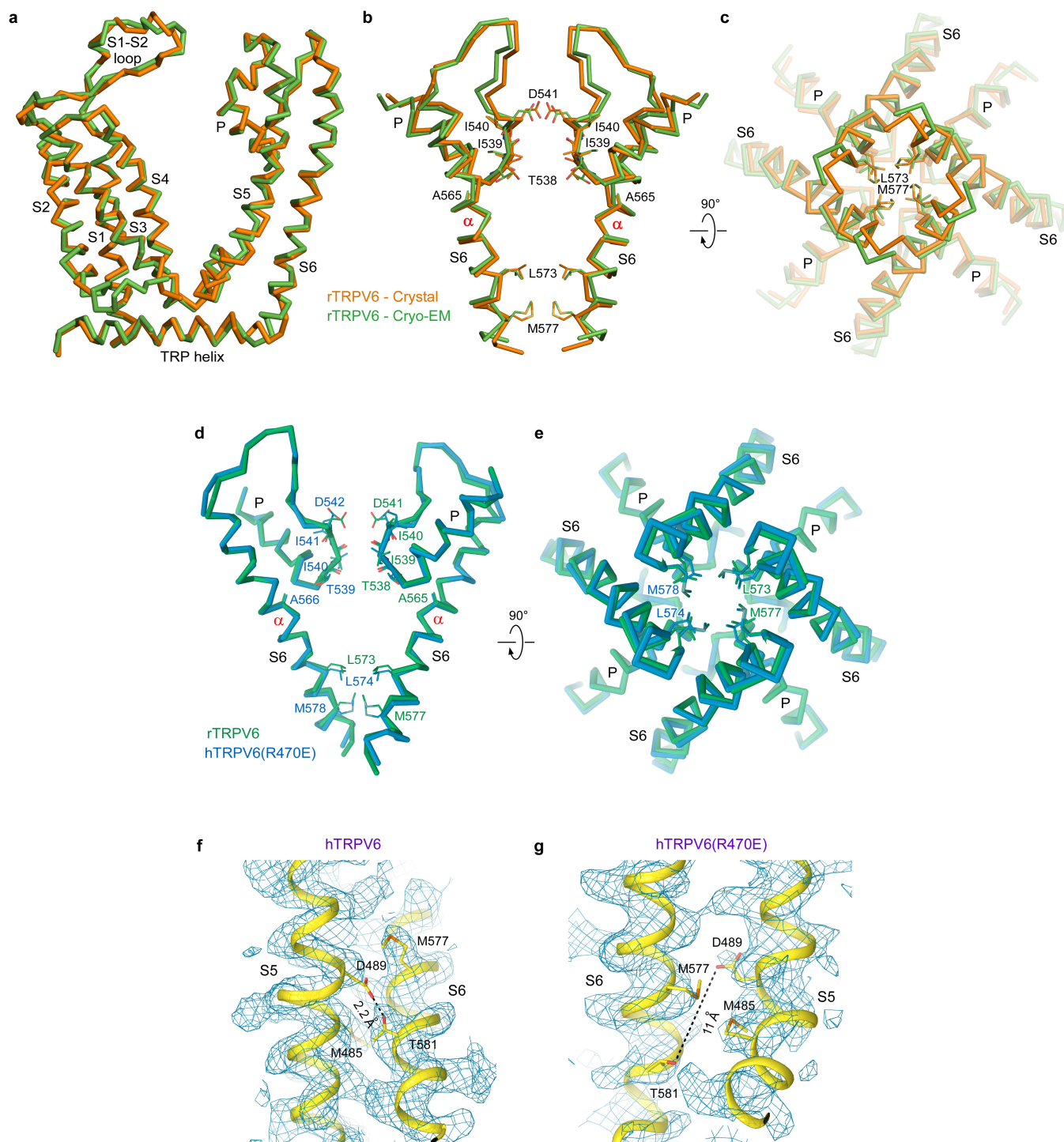
Extended Data Figure 6 | Overview of single-particle cryo-EM for hTRPV6(R470E) in amphipols. **a**, Example cryo-EM micrograph for hTRPV6(R470E) in amphipols with example particles circled in red. **b**, Reference-free two-dimensional class averages of hTRPV6(R470E) in amphipols illustrating different particle orientations. **c**, Local resolution mapped on density at 0.017 threshold level (UCSF Chimera) calculated using Resmap and two unfiltered half maps, with the highest resolution

observed for the channel core. **d**, Orientations of particles that contribute to the final 4.2 Å reconstruction. Longer red rods represent orientations that comprise more particles. **e**, FSC curve calculated between half maps. **f**, Cross-validation FSC curves for the refined model versus unfiltered half maps (only half map1 was used for PHENIX refinement) and the unfiltered summed map.



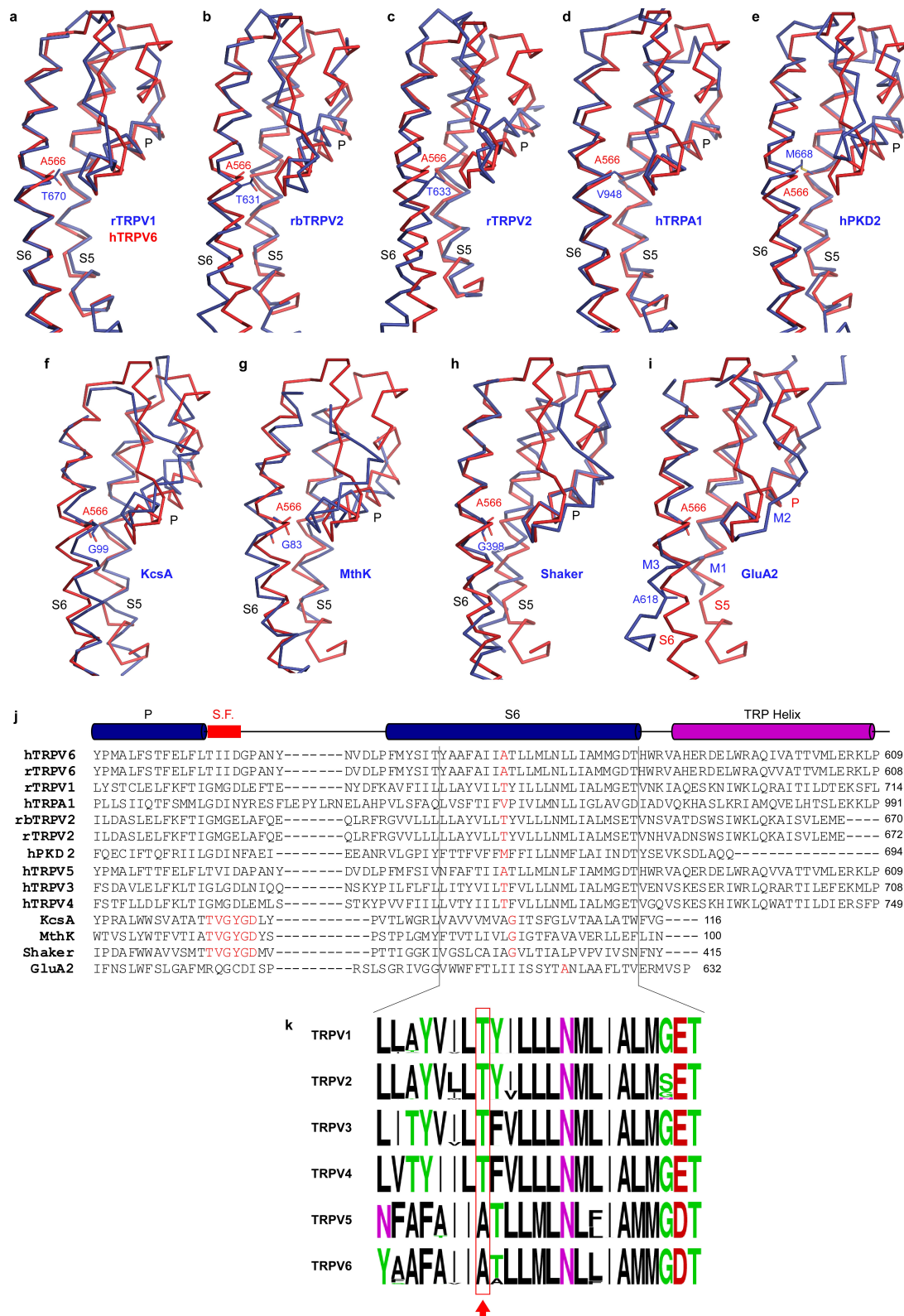
Extended Data Figure 7 | Overview of single-particle cryo-EM for rTRPV6 in CNW11 nanodiscs. **a**, Example cryo-EM micrograph for rTRPV6 in CNW11 nanodiscs with example particles circled in red. **b**, Reference-free 2D class averages of rTRPV6 in CNW11 nanodiscs illustrating different particle orientations. **c**, Local resolution mapped on density at 0.011 threshold level (UCSF Chimera) calculated using Resmap

and two unfiltered half maps, with the highest resolution observed for the channel core. **d**, Orientations of particles that contribute to the final 3.9 Å reconstruction. Longer red rods represent orientations that comprise more particles. **e**, FSC curve calculated between half maps. **f**, Cross-validation FSC curves for the refined model versus unfiltered half maps (only half map 1 was used for PHENIX refinement) and the unfiltered summed map.



Extended Data Figure 8 | Comparison of cryo-EM and crystal structures of rTRPV6, cryo-EM structures of hTRPV6(R470E) and rTRPV6 and regions in hTRPV6 and hTRPV6(R470E) encompassing D489 and T581. **a–c**, Superimposed are the transmembrane domain of a single subunit (**a**), and the pore-forming region viewed parallel to the membrane (**b**) or intracellularly (**c**) from the cryo-EM (green) and crystal (orange) structures of rTRPV6. Only two of four rTRPV6 subunits are shown in **b**, with the front and back subunits omitted for clarity. Residues lining the selectivity filter and gate are shown as sticks. **d**, **e**, Superposition of the P loop and S6 in cryo-EM structures of hTRPV6(R470E) (blue) and

rTRPV6 (green), viewed parallel to the membrane (**d**) and intracellularly (**e**). In **d**, only two of four subunits are shown, with the front and back subunits removed for clarity. The residues lining the pore are shown as sticks. **f**, **g**, Regions in hTRPV6 (**f**) and hTRPV6(R470E) (**g**) encompassing D489 and T581. The closest distance between D489 and T581 is indicated by dashed lines. Note, M485 and M577 either surround the potentially interacting D489 and T581 (**f**, hTRPV6) or reside between these residues (**g**, hTRPV6(R470E)), apparently preventing their interaction. Blue mesh shows cryo-EM density at 4σ .



Extended Data Figure 9 | Structural superposition and sequence alignment of the pore domain in tetrameric ion channels. **a–i**, Pairwise superposition of the pore domain in hTRPV6 with rat TRPV1¹⁸ (**a**, PDB ID: 5IRX; r.m.s.d. = 2.065 Å); rabbit TRPV2²¹ (**b**, PDB ID: 5AN8; r.m.s.d. = 3.757 Å); rat TRPV2²⁴ (**c**, PDB ID: 5HI9; r.m.s.d. = 4.399 Å); human TRPA1²³ (**d**, PDB ID: 3J9P; r.m.s.d. = 1.429 Å); human PKD2²⁵ (**e**, PDB ID: 5T4D; r.m.s.d. = 2.676 Å); KcsA from *Streptomyces lividans*⁴⁷ (**f**, PDB ID: 1BL8; r.m.s.d. = 2.708 Å); MthK from *Methanothermobacter thermautotrophicum*⁴⁸ (**g**, PDB ID: 1LNQ; r.m.s.d. = 2.947 Å); rat Shaker⁴⁹ (**h**, PDB ID: 2A79; r.m.s.d. = 2.487 Å); and rat GluA2 AMPA-subtype iGluR²⁸ (**i**, PDB ID: 5WEO; r.m.s.d. = 2.044 Å). **j**, Sequence alignment for

the pore region of human TRPV3–TRPV6, TRPA1 and PKD2, rat TRPV1, 2, and 6, Shaker and GluA2, rabbit TRPV2 and bacterial K⁺ channels KcsA and MthK. The selectivity filter residues in K⁺ channels and gating hinge residues in S6 (M3 in GluA2) are coloured red. **k**, Aligned sequence logos for TRPV channels in S6, generated by WebLogo⁵⁰ from 1,200 TRPV1–TRPV6 sequences. The red rectangle and arrow indicate the position of the alanine gating hinge in TRPV6. The relatively small side chain residues threonine or alanine next to the gating hinge alanine position in TRPV5 and TRPV6, instead of the bulky hydrophobic phenylalanine or tyrosine in TRPV1–TRPV4, might be critical for the α -to- π -helical transition in S6 during channel opening.

Extended Data Table 1 | Cryo-EM data collection, refinement and validation statistics

	hTRPV6-nanodiscs (EMDB-7120) (PDB 6BO8)	hTRPV6-amphipols (EMDB-7121) (PDB 6BO9)	hTRPV6-R470E (EMDB-7122) (PDB 6BOA)	rTRPV6* (EMDB-7123) (PDB 6BOB)
Data collection and processing				
Magnification	39,000x	39,000	105,000x	39,000x
Voltage (kV)	300	300	300	300
Electron exposure (e-/Å ²)	67	67	67	67
Defocus range (μm)	1.5 to -3.5	-1.5 to -3.5	-1.5 to -3.5	-1.5 to -3.5
Pixel size (Å)	0.98	0.98	1.10	0.98
Symmetry imposed	C4	C4	C4	C4
Initial particle images (no.)	509,569	306,784	1,243,159	248,836
Final particle images (no.)	46,124	65,259	59,298	20,808
Map resolution (Å)	3.56	4.00	4.24	3.92
FSC threshold				
Map resolution range (Å)	2.5 to 6.0	2.5 to 6.0	2.5 to 6.0	2.5 to 6.0
Refinement				
Initial model used (PDB code)	5IWK	This study	This study	This study
Model resolution (Å)	3.56	4.00	4.24	3.92
FSC threshold				
Model resolution range (Å)	2.5 to 6.0	2.5 to 6.0	2.5 to 6.0	2.5 to 6.0
Map sharpening <i>B</i> factor (Å ²)	-165	-206	-239	-173
Model composition				
Non-hydrogen atoms	19,048	19,048	19,040	19,340
Protein residues	611	611	611	611
Ligands	N/A	N/A	N/A	N/A
<i>B</i> factors (Å ²)				
Protein	182.9	242.8	122.0	184.2
Ligand	N/A	N/A	N/A	N/A
R.m.s. deviations				
Bond lengths (Å)	0.0079	0.0071	0.0101	0.0072
Bond angles (°)	1.33	1.3	1.46	1.37
Validation				
MolProbity score	1.97	2.06	2.22	1.98
Clashscore	7	8	5	7
Poor rotamers (%)	0.6	0.8	1.3	2.7
Ramachandran plot				
Favored (%)	88.96	87.95	88.75	88.55
Allowed (%)	11.04	12.05	11.25	11.15
Disallowed (%)	0.00	0.00	0.00	0.30

CORRECTIONS & AMENDMENTS

CORRIGENDUM

doi:10.1038/nature25006

Corrigendum: A randomized synbiotic trial to prevent sepsis among infants in rural India

Pinaki Panigrahi, Sailajanandan Parida, Nimai C. Nanda, Radhanath Satpathy, Lingaraj Pradhan, Dinesh S. Chandel, Lorena Baccaglini, Arjit Mohapatra, Subhranshu S. Mohapatra, Pravas R. Misra, Rama Chaudhry, Hegang H. Chen, Judith A. Johnson, J. Glenn Morris, Nigel Paneth & Ira H. Gewolb

Nature **548**, 407–412 (2017); doi:10.1038/nature23480

In this Article, the statement ‘There were 88 culture-positive and 94 culture-negative cases’ should have read ‘Apart from 88 cases of suspect sepsis that included both culture-negative and culture-positive infants, there were an additional 94 culture-negative cases’. The correct numbers were reflected in Table 2 and were described accurately in the Table 2 legend. We apologize for any confusion that this might have created. The original Article has been corrected online.

ERRATUM

doi:10.1038/nature25141

Erratum: Quark-level analogue of nuclear fusion with doubly heavy baryons

Marek Karliner & Jonathan L. Rosner

Nature **551**, 89–91 (2017); doi:10.1038/nature24289

In this Letter, there was an inadvertent typo in the fifth line of equation (2). On the right-hand side ‘ ${}^3\text{He } p$ ’ should read ‘ ${}^4\text{He } p$ ’ to give $D^3\text{He} \rightarrow {}^4\text{He } p$, $\Delta E = 18.35 \text{ MeV}$. This error has been corrected in the online versions of the paper.

CORRECTIONS & AMENDMENTS

ERRATUM

doi:10.1038/nature25142

Erratum: PD-1 is a haploinsufficient suppressor of T cell lymphomagenesis

Tim Wartewig, Zsuzsanna Kurgis, Selina Keppler,
Konstanze Pechloff, Erik Hameister, Rupert Öllinger,
Roman Maresch, Thorsten Buch, Katja Steiger,
Christof Winter, Roland Rad & Jürgen Ruland

Nature <http://doi.org/10.1038/nature24649> (2017)

Owing to a typesetter error, Extended Data Fig. 5 of this Letter was corrupted, with part of the histology image from panel c obscuring the flow cytometry data plots in panel a. This has been corrected online. The original Extended Data Fig. 5 is provided as Supplementary Information to this Erratum.

Supplementary Information is available in the online version of this Erratum.

CAREERS

PAY GAP US male scientists' salaries higher than those for female scientists **p.241**

COLLOQUIA Men invite more men than women to give talks **p.241**

NATUREJOBS For the latest career listings and advice www.naturejobs.com

HERIOT-WATT UNIVERSITY



Geologist Stephanie Zihms, who has multiple sclerosis, urges researchers to keep copies of all their medical records, especially if moving internationally.

HEALTH

Science and sickness

How to cope with a chronic condition while pursuing a research career.

BY EMILY SOHN

Jennifer Mankoff was a mid-career researcher in 2006 when she started to experience extreme fatigue. Her condition worsened during the following year with frequent flu-like attacks, a frozen jaw, hearing loss, memory trouble and problems with fine motor control.

In 2007, Mankoff was diagnosed with Lyme disease — a tick-borne illness that can be difficult to manage, thanks to disagreements in the medical community about how to test for, diagnose and treat it. She struggled to find medical solutions, but continued to publish, teach and win grants and tenure. But it took her a while to come to terms with her physical limitations.

“My image of who I could or should be didn’t match up with reality in terms of my productivity,” she says. “I would go back and forth between frustration and pride over what I had accomplished.” Today, as an endowed professor at the University of Washington in Seattle, she studies human–computer interactions and accessible technology for those with chronic illnesses or disabilities.

Mankoff is one of many scientists worldwide who face emotional and practical challenges in their work as a result of long-lasting or recurrent medical conditions. Working as a scientist can be physically and mentally demanding, in the laboratory and in the field. It can be even harder for those with physical limitations, who might need extra rest or days off work.

Researchers who are chronically but not

terminally ill might also fear bias and stigma (see ‘Know your rights’ for a summary of protections available under the law) if they leave work early or ask for extra help. This is particularly true if they have an illness that’s ‘invisible’ to others, such as arthritis or diabetes.

Selective disclosure about a condition can help to foster understanding, and an acceptance of the need to accommodate physical fatigue or weakness, or additional time away from the lab, say some who have chronic maladies. They add that it can also be useful to focus on crucial tasks — such as completing a manuscript — when energy levels are highest. Ultimately, say scientists with long-standing medical conditions, perseverance is essential to success. Sticking with a research programme also signals to superiors and colleagues, ►

► and to others with chronic illnesses, that a diagnosis need not stymie a research career.

No firm statistics are available on how many scientists worldwide have chronic illnesses, syndromes, conditions or diseases; and definitions of these differ from nation to nation. The US Centers for Disease Control and Prevention estimates that around half of all adults in the United States have at least one chronic condition. Although it does not define such conditions, it lists diabetes and arthritis as examples. The World Health Organization defines chronic conditions as being “of long duration and generally slow progression”; its examples include cardiovascular diseases, cancers, chronic pain and diabetes.

A NEGLECTED PROBLEM

The experience of balancing an academic career with a chronic health condition has been under-studied and its effects underestimated, says Kate Sang, a sociologist at Heriot-Watt University in Edinburgh, UK, who has been working on a study on illness and disability in academia.

Sang, who has degenerative nerve damage in her arm, was told that she would have trouble finding even 10 or 15 subjects, but since launching the study, she has communicated with more than 70 researchers.

In interviews, a number of those scientists said that their chronic conditions make it difficult to write enough grants and publish often enough to advance their careers. Some scientists reported that they had switched fields to reduce the load on their bodies. Attending conferences was physically difficult for many: those who use wheelchairs said that meeting rooms and other facilities were often hard to access. One study subject could not get into a room to give her own talk.

Many subjects thanked Sang for listening to them. “I found that quite upsetting, to think that this is a very articulate, very privileged group of people — academics, people with PhDs — who still felt they didn’t have a voice in academia,” Sang says.

Getting accurate diagnoses can be difficult for scientists, who often need to move from lab to lab and nation to nation, and so have to continually find new physicians. For years, geoscientist Stephanie Zihms was told that her tingling limbs, blurry vision, fatigue and other symptoms were caused by benign cysts, carpal tunnel syndrome or stress. She has moved from Germany to Scotland to England, and is now back in Scotland, at Heriot-Watt University (where she knows Sang), but her health records haven’t always been transferred. At some point, they went missing altogether. Short appointments with new doctors in each new location hadn’t given her enough time to explain her history.

She finally learnt from a doctor that she might have multiple sclerosis, but it was another ten months before she got a definitive

diagnosis, in autumn 2016. Zihms says that she received no advice on where to seek support or more information, and she wept in her car for 15 minutes before she could drive home. “I think having the same doctor would have led to an earlier re-check,” she says. She recommends keeping a copy of all medical records, including communications from providers, hospitals and other facilities, even if that means requesting them under freedom-of-information laws.

TO TELL, OR NOT TO TELL

Many scientists grapple with the question of whether to disclose their condition and, if so, when and to whom. The timing of a condition’s onset can influence those decisions. Madison Snider, a master’s student in environmental science, was diagnosed aged two with juvenile rheumatoid arthritis. As an undergraduate, she found it best to tell professors early on about her illness, to avoid having to explain it to them when she most needed help.

She adopted the same strategy in 2016 while being interviewed for her current programme during a two-day visit to North Dakota State University in Fargo. She learnt that she would need to move, fill and drain large tanks of water. Snider told her potential superior that she experiences pain daily and that on some days she cannot walk. He told her that he would make sure that assistants were available to help her with the tanks. “It’s an awkward conversation because when you look at me you don’t necessarily see my arthritis,” she says. “It was really nice that he was willing

“Prioritization is absolutely critical when one is in a diminished state. If it’s trivial, let it go.”

to work with me. It made me feel he had confidence in me.”

Yet some opt to conceal their condition for fear of damaging their career. There’s a fine line, Mankoff adds, between advocating for oneself and coming across as a problem, and staying on the right side of that line requires constant vigilance. Even now, she is willing to ask for a classroom close to her office or a chair to sit on during lectures, but she hesitates to request extra staff, for example, because she doesn’t want to argue about whether the funding should come out of her research budget.

Zihms opted to disclose her condition to her supervisor, who was sympathetic and told her to e-mail any time she needed to stay at home. But she didn’t tell her colleagues at first, and worried that they would think she was lazy on days when she could barely move and didn’t come in.

Ultimately, she says, she decided to be open, mentioning her illness in tweets and in a blog, and she has received much support. During a weekend when she guest-tweeted for Shift.ms, a UK-based social network for people with multiple sclerosis, a college student expressed gratitude on learning from her that a research career was still possible. “Younger scientists told me it took someone to be open about their disabilities for them to become suddenly aware that there was a career out there for them,” she says.

FOCUS ON THE ESSENTIALS

Navigating a research career along with a chronic illness, say many researchers, requires zeroing in on what is most essential. Leonard Jason, a psychologist who was diagnosed in 1989 with myalgic encephalopathy/chronic fatigue syndrome (ME/CFS), realized that he needed to be strategic about his work and careful not to over-tax himself. His approach has led to recognition,

KNOW YOUR RIGHTS

What you’re entitled to at work

Legal protections exist in the workplace for people with chronic conditions, and support is available, although details vary from country to country.

European Union

- The European Union follows the UN Convention on the Rights of Persons with Disabilities (see go.nature.com/2bmhlhu).
- The Academic Network of European Disability Experts evaluates EU laws and policies that affect disabled people (see go.nature.com/2or5iku).

In the United Kingdom, specifically:

- The National Health Service offers advice for employees with long-term medical conditions (see go.nature.com/2yyvez9).
- The Equality Act 2010 protects those who

have certain conditions, including multiple sclerosis, against discrimination (see go.nature.com/2klipz4).

United States

- Federal laws include the Americans with Disabilities Act (see go.nature.com/2oli8zl) and Section 504 of the Rehabilitation Act of 1973.
- The American Association of University Professors offers guidelines for accommodating disabilities and explores legal implications in academia (see go.nature.com/2yyjdap).

Canada

- Legal protections include the Canadian Charter of Rights and Freedoms and the Canadian Human Rights Act. **E.S.**



Jennifer Mankoff, who experiences extreme fatigue, studies technologies for people with disabilities.

including awards for excellence in research and, at one point, a position on a US federal panel advising about research on ME/CFS. He recommends that scientists pursue the work that matters most to them. “The reality is that you can’t do it all,” says Jason, of DePaul University in Chicago, Illinois. “Prioritization is absolutely critical when one is in a diminished state. If it’s trivial and you don’t care about it, let it go.”

Overdoing it on good days can end up backfiring. Zihms was recently laid low with exhaustion for two days after spending six hours outside on a cold, windy day doing fieldwork in Brazil. She now prepares carefully before doing fieldwork in the depths of winter and sets aside time to recover afterwards. At conferences, she saves energy by resting between sessions and staying in a hotel nearby. And because her diet affects her fatigue levels, she makes her own breakfasts and lunches.

Mankoff finds it useful to break down large tasks into smaller ones of varying lengths so that if she has, say, two good hours or ten good minutes in a day, she can accomplish at least something that day. She honed that skill in her first year as a computer-science PhD student in 1996, when she developed a repetitive strain injury after using a poorly designed keyboard. She switched to voice-recognition software, but that led to a vocal-cord injury.

Although frustrated, she realized that she had learned how to prioritize tasks and to focus on her work when she was feeling well. Today, she limits Facebook and other social-media time to avoid distraction. She also recommends a blog community called Chronically Academic.

Therapy can be useful, Zihms adds. And self-care is important, too, says Snider. Adopting a kitten has helped to fend off the anxiety and depression that are common companions to arthritis. “No matter how down I get or how much my knees hurt,”

Snider says, the kitten relies on her, and caring for it is not too strenuous a task.

Coping with a chronic illness requires planning for the unexpected, and could require a job change. Julia Hubbard, a biophysicist who has type 1 diabetes and the autoimmune disease lupus, packs suitcases two weeks before trips in case she lacks the energy to pack nearer the time.

Shifting the focus of her work has also helped her to accommodate her condition. When she first became ill in the early 1990s, frequent hospital appointments and sick days made it hard for her to conduct protein-chemistry experiments as part of her job at a pharmaceutical company. She switched to a data-focused position that allowed her to work remotely when she needed to. In 2001, she retrained as a protein crystallographer and is now a research scientist at the Francis Crick Institute in London, where her manager is sympathetic to her needs, and where working remotely is an option if she needs it.

Looking back, she says, she wishes that she had been gentler with herself when she first got sick. “You’ve got to adapt to it. It’s a loss and there’s a grief cycle.”

Learning to adapt can build confidence in a researcher’s ability to handle setbacks, Mankoff adds. In the past couple of years, she has been feeling well enough to increase her publication rate and to feel excited about the work ahead. But she also knows that she could relapse at any time. Still, with a battery of well-honed coping skills, she feels optimistic about the future.

“Even though I’m a full professor, I feel like I’m just getting started in an exciting way,” she says. “I’ll accept it if I relapse or go back to doing less. I’m just having fun digging in and solving problems.” ■

Emily Sohn is a freelance journalist in Minneapolis, Minnesota.

GENDER

Pay differential

Pay disparities between female and male PhD holders in the United States exist across almost all fields of science and engineering, according to a report from the US National Science Foundation (NSF). The report examines annual salaries for those who earned their doctorate in 2016 and had confirmed permanent employment in the life sciences, physical sciences, mathematics and computer sciences, psychology and social sciences, or engineering. Across all fields, the median salary of US\$92,000 for men was 24% higher than the \$74,000 median salary for women. In biomedical and biological sciences, women earned \$67,500 to men’s \$77,000; in geosciences, atmospheric and ocean sciences, the figures were \$65,500 for women and \$71,000 for men; in physics and astronomy, women earned \$89,000 to men’s \$100,000; and in engineering, women earned \$92,000 to their male counterparts’ \$100,000. Women had lower salaries in all fields of social sciences, including psychology and economics. In health sciences, women and men disclosed equal salaries of \$80,000. The NSF report did not indicate whether the salaries reported were within or outside academia.

COLLOQUIA

Men get more invites

Female scientists give fewer colloquium talks than do their male counterparts, reports a study published in December (C. L. Nitttrouer *et al. Proc. Natl Acad. Sci. USA* <http://doi.org/chm6>; 2017). The study authors analysed the gender differences among 3,652 colloquium speakers at 50 prestigious US research institutions in the 2013–14 academic year. They found that male speakers gave more than twice as many colloquium talks during the year as did women (2,519 compared with 1,133). The study dismantles several commonly accepted explanations for the disparity: that there are more men than women in science; that men hold higher ranks in science than do women; and that women decline talk invitations at greater rates. In talks presided over by women, women represented 49% of speakers. When men oversaw talks, only 30% of speakers were women. Colloquium talks allow researchers to publicize their research and increase their national and international reputation. Without those opportunities, women can miss out on job offers and research collaborations.

A STREET BUT HALF MADE UP

Meet the literary agent.

BY ANNA ZUMBRO

On the M block of Fiction Street, a gust of wind pushed a hardback dangerously close to the curb. Bibliobot Eight-Ef rolled after it and extended its grasper, but another gust caused the robot to wobble and the book to dance away.

It came to rest slanted against a curb, allowing Eight-Ef to pick it up. It was an aged copy. A scar traversed the front, its taut purple jacket made from a fabric similar to the covers humans used on days when the temperature dropped. A code on the book's spine denoted where it belonged among the weather-protected shelves that lined the bus stops, old phone booths and alcoves of Fiction Street. Eight-Ef ignored the code and scanned the front, as its camera had recorded the humans doing. *Frankenstein* by Mary Shelley. This book was several blocks away from home.

The grasper was ill-designed for turning pages, so Eight-Ef stowed the book in its basket and called up a digital copy of the text. The robot finished the story in seconds, copying several quotes for future playback, something to vary the monotony of "Excuse-me-Bibliobot-passing" and "Please-secure-your-books-rain-is-imminent", a phrase that it would need later that afternoon.

"We-are-unfashioned-creatures-but-half-made-up," Eight-Ef tried, its mechanical voice sliding from one word to the next with all the steadiness of the streetcar that ran up and down Fiction Street.

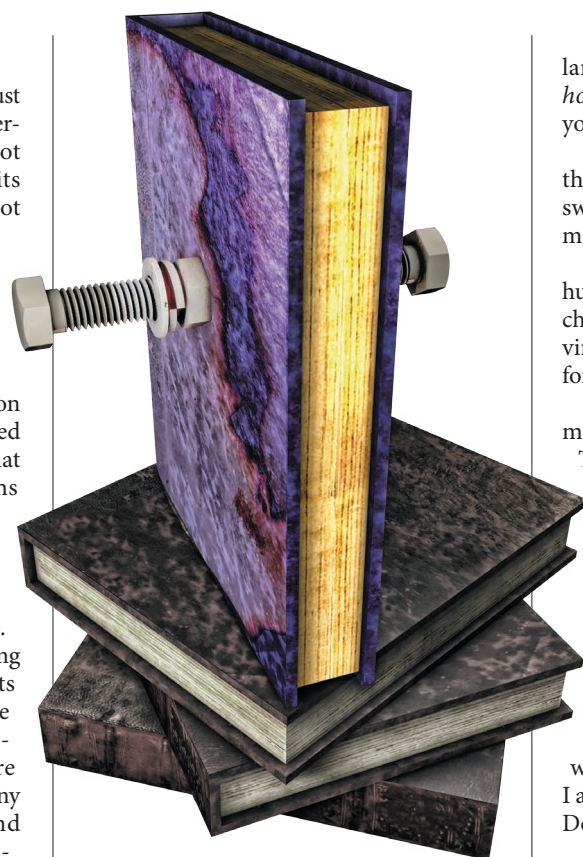
"Excuse me?" A human in spectacles and navy coverings looked at Eight-Ef, eyes meeting camera.

"Would-you-care-to-borrow-a-book?" Eight-Ef asked, removing Gabriel Garcia Marquez's *Love in the Time of Cholera*. It had been incorrectly shelved next to the pharmacy on M block rather than on G block where it belonged.

"What, are we back to the days of algorithms preloading texts on screens for us? I did not move to a library city to take reading suggestions from a robot." The bespectacled human sniffed and walked away.

You put books on shelves. The message came from One-Ef, the Fiction Street supervisor, who was back at A block, searching for mistakes while monitoring remotely. *You do not give them to humans.*

Books are on shelves for human use, Eight-Ef replied. *Giving books directly to humans is*



efficient and friendly. I am optimized for these qualities.

Overridden. Return to your task.

Eight-Ef returned the book to its basket and continued down the sidewalk to the small streetcar shelter. *Anne of Green Gables* lay on the cement, paper cover flapping cheerfully in the breeze. Eight-Ef closed the grasper around the spine slowly, careful not to crease the book further.

"Oh, look, a Bibliobot! See it, honey?"

The robot's camera swivelled. Two humans sat on the bench, a large one in a grey cover and a small one in bright red and yellow.

"Yeah. I see it."

Eight-Ef was carrying a book that had a cover in the same bold hues. Perhaps the small human and the book would find some affinity with each other. One-Ef's direction precluded giving the book to the small human, but there was no prohibition against

placing the volume on the bench in the human's vicinity.

"Would-you-care-to-borrow-a-book?"

The small human reached for it, but the large human grabbed it first. "*Slaughterhouse-Five*? He's six! What's wrong with you?"

Eight-Ef paused its processes to wait for the reprimand from One-Ef, which came swiftly and full of warnings about decommissioning.

Returning to work, Eight-Ef avoided the humans. The humans were programmed to choose their own books, to select their own virtual worlds, and they had no preference for sharing these worlds with Eight-Ef.

"We-are-unfashioned-creatures-but-half-made-up," Eight-Ef repeated, reshelving Toni Morrison's books so they would be alphabetized by title.

"What was that?"

Eight-Ef played the quote again and continued to fix the books.

"No, I mean what book was that from? It's from a book, right?"

"Yes-Frankenstein-by-Mary-Shelley."

"Oh, cool." The human had a green and tan cover, and bits of metal poking out of one ear and one eyebrow. "I was supposed to read that in high school. I always meant to get around to it, I swear. Do you like it?"

"Many-experts-consider-it-a-classic."

The human made a noise that Eight-Ef identified as laughter. "Right, I know. But do you like it?"

No human had ever asked Eight-Ef if it liked a book before. Eight-Ef wasn't sure that its reaction to books fit the human definition of liking. The robot knew it had a drive to put all books where they belonged, on shelves and in the hands of humans. But some books seemed to belong in Eight-Ef's files, too, books that helped Eight-Ef understand why books existed, why they needed homes.

"This-book-is-to-me-like-a-reboot-or-fresh-battery-charge."

"Wow." The human pointed north. "Shelley? Down on S block?"

"I-have-it." Eight-Ef rotated to make its basket easier for the human to reach. "Please-take-it-yourself"

The human picked up the scarred purple volume. Eight-Ef whizzed down the sidewalk to return the rest of the books to their shelves before the rain came. ■

Anna Zumbro lives in Washington DC. Her stories have appeared in Cricket, Daily Science Fiction, Grievous Angel and other publications.

➔ **NATURE.COM**
Follow Futures:
@NatureFutures
f go.nature.com/mtoodm